1 **Trajectories of change in Mediterranean Holocene vegetation through classification of**

2 **pollen data**

3 Ralph M. Fyfe[1], Jessie Woodbridge and C. Neil Roberts

4 School of Geography, Earth and Environmental Sciences, University of Plymouth, Plymouth,

5 PL4 8AA, UK

6 [1] corresponding author: ralph.fyfe@plymouth.ac.uk; +44 1752 585929.

7 Orchid ID 0000-0002-5676-008X

8

9 **Abstract**

10 Quantification of vegetation cover from pollen analysis has been a goal of palynologists since

11 the advent of the method in 1916 by the great Lennart von Post. Pollen-based research

12 projects are becoming increasingly ambitious in scale, and the emergence of spatially

13 extensive open-access datasets, advanced methods and computer power has facilitated sub-

14 continental analysis of Holocene pollen data. This paper presents results of one such study,

15 focussing on the Mediterranean basin. Pollen data from 105 fossil sequences have been

16 extracted from the European Pollen database, harmonised by both taxonomy and

17 chronologies, and subjected to a hierarchical agglomerative clustering method to synthesise

18 the dataset into 16 main groupings. A particular focus of analysis was to describe the

19 common transitions from one group to another to understand pathways of Holocene

20 vegetation change in the Mediterranean. Two pollen-based indices of human impact (OJC:

21 Oleaceae, *Juglans, Castanea*; API: anthropogenic pollen indicators) have been used to infer

22 the degree of human modification of vegetation within each pollen grouping. Pollen-inferred

23 cluster groups that are interpreted as representing more natural vegetation states show a

24 restricted number of pathways of change. A set of cluster groups were identified that closely

25 resemble anthropogenically-disturbed vegetation, and might be considered anthromes

26 (anthopogenic biomes). These clusters show a very wide set of potential pathways, implying

27 that all potential vegetation communities identified through this analysis have been altered in

28 response to land exploitation and transformation by human societies in combination with

29 other factors, such as climatic change. Future work to explain these ecosystem pathways will

30 require developing complementary datasets from the social sciences and humanities

31 (archaeology and historical sources), along with synthesis of the climatic records from the

32 region.

33

Keywords

35 Pollen, classification, cluster analysis, Mediterranean, Holocene, change analysis

Acknowledgements

47

Introduction

48

49   Pollen analysis allows past vegetation to be described, with the signal mediated through

50   taphonomic and depositional processes. A wide variety of methods have been established to

51   translate pollen data into meaningful vegetation units, including biomisation (Prentice et al.

52   1996), pseudobiomisation (Fyfe et al. 2010; Woodbridge et al. 2014), modern analogue

53   methods for identifying modern equivalents to fossil assemblages (e.g. Gaillard et al. 1994)

54   and quantifying vegetation cover through model-based transformation (Sugita 2007a, b;

55   Gaillard et al. 2010; Trondman et al. 2015). Understanding the processes that lead to changes

56   in vegetation is a next logical step after palaeo-vegetation classification and description.

57   Much effort is focussed on understanding the role of anthropogenic forcing in transformation

58   of vegetation, something that von Post (1946) largely overlooked, in spite of his engagement

59   with archaeological research programmes.

60   Humans have transformed natural environments over many millennia, from Neolithic

61   farming to modern agriculture. Williams et al. (2015) have argued that technological

62   development has rapidly transformed ecological structure and dynamics, including vegetation

63   (impacting on species richness, evenness, and biomass), resulting in changes from semi-

64   natural or natural systems to human-modified vegetation. Ellis and Ramankutty (2008)

65   describe the transformation from 'natural' to 'anthropogenic' biomes, using the term

66   'anthromes'. The development of 'anthromes' likely occurred over many millennia (Ellis et

67   al. 2010; Ellis 2011; Ruddiman et al. 2016), but transformation of vegetation at the global

68   scale is not recognised until recent centuries (Ellis 2011). It is possible to identify anthromes

69   at a regional (sub-continental) scale, particularly in areas with longer histories of complex

70   societies. Understanding the development of (and trajectories towards) anthromes can be

71   facilitated by meta-analysis of pollen data across regions and continents (e.g. Fyfe et al. 2015;

72   Trondman et al. 2015, although neither explicitly describe anthromes). The emergence of

73   such novel ecosystems has been a focus of interest for conservation management, but, as

74   Perring and Ellis (2013) have argued, novel ecosystems are frequently ancient rather than

75   recent developments and ecosystem novelty also depends on when the reference baseline is

76   set (Radeloff et al. 2015). In the Mediterranean, sedentary village life began by 8000 BP and

77   agriculture was established across the majority of the European Mediterranean region by

78   6000 BP (Roberts 2013). By the 4th millennium BP, complex societies operated within a

79   clearly established 'world' system (Butzer 2005), periodically punctuated by collapses of

80   higher-order socio-political structures. In addition, fire has long been used as a tool of

81   environmental management, notably to encourage re-growth of new vegetation for livestock

82   grazing and browsing (Vannière et al. 2010).

83   The divergence of "natural" from modern vegetation in the Mediterranean has previously

84   been explored by Collins et al. (2012), by comparison of fossil assemblages from 6000 BP

85   with recent pollen samples from the same sites, and in the earlier BIOME6k project (Prentice

86   et al. 2000; Roberts et al. 2004). Clear differences are described between 6000 BP as a

87   reference baseline, and modern time periods, most notably via increasing amounts of open

88   ground, and the establishment and spread of disturbed and anthropogenically-modified

89   vegetation. This paper seeks to understand transitions from one vegetation state to another

90   using the example of the Mediterranean. It will assess the extent to which

91   pathways/transitions from one state to another have been predictable and replicated between

92   sites. The methodological basis, which is described in greater detail below, involves

93   "unsupervised" data classification, rather than imposed quantification (c.f. biomisation:

94   Prentice et al., 1996), to produce a taxonomy of pollen samples, grouped by similarity in their

95   assemblages. The paper draws on a large number of fossil pollen sequences from across the

96   Mediterranean basin. Detailed analysis that compares the unsupervised classification of these

97   pollen data to pre-existing classifications (e.g. biomisation) and modern vegetation mappings

98   is described in Woodbridge et al. (in review). Recent work by Felde et al. (2014, 2016) has

99   demonstrated the clear potential of such an approach in the analysis of large pollen datasets,

100   something that even a decade ago would not have been possible without the advent of high-

101   performance computing, and the establishment of clear statistical approaches for the analysis

102   of palaeoecological datasets (Birks et al. 2012).

103

Materials and methods

Modern and fossil pollen datasets

106   The approach to classification of samples used here is Ward's hierarchical agglomerative

107   clustering method (Ward 1963; Murtagh and Legendre 2014), using Euclidean distance as the

108   dissimilarity measure between samples. Whilst a large number of alternative methods exist

109   for the classification of palaeobotanical data, careful experimentation has shown that different

110   methods commonly produce similar results (Felde et al. 2014, 2016). As the purpose here is

111   to assess trajectories of change between classes of pollen assemblages, the simpler Ward's

112   method has been employed over more complex approaches; Woodbridge et al. (in review)

113   discuss some alternative approaches to classification of the same dataset. An advantage of a

114   hierarchical approach to clustering data is also the ability to (dis)aggregate at different levels

115   across the dendrogram (Garcia-Madrid et al. 2014). Analyses were undertaken using the R

116   package VEGAN (Oksanen et al. 2016).

117   All pollen data used have been extracted from open-access online data stores. Sub-fossil

118   pollen count data was obtained from the European Pollen Database (EPD: May 2016 version:

119   Leydet 2007-2017). The EPD is a community-driven database that archives site metadata,

120   chronological information and raw count data (Fyfe et al. 2009). Pollen sites located in areas

121   with characteristic Mediterranean climate and vegetation have been selected (Fig. 1), using

122   definitions from Ozendaa and Borela (2000) and Izdebski et al. (2015). Figure 1B and 1C

123   show that Mediterranean climate space is well represented in both our modern and fossil

124   pollen site data. Each fossil sample extracted from the EPD was assigned a calibrated age

125   estimate on the basis of established core chronologies (Giesecke et al. 2014; Leydet 2007-

126   2017). The pollen count data from each site has been summed into contiguous 200-year time

127   windows from 11000-10800 to 200- -65 cal BP (with AD1950 as present day). Additionally,

128   'modern' pollen count data was taken from the European Modern Pollen Database (EMPD:

129   Davis et al. 2013), a repository of surface pollen count data, including date of collection,

130   locational information and additional site metadata. The taxonomy of the EPD and EMPD

131   was harmonised and simplified, taking the EPD as the base. Where possible a detailed level

132   of taxonomic resolution was retained, particularly for key indicator species of disturbance

133   (e.g. ribwort plantain: *Plantago lanceolata*); other taxa were combined at a higher taxonomic

134   level owing to differences in recording between analysts across Europe (e.g. all evergreen

135   oak species and varieties). All taxa below 1% of the pollen sum that occurred in fewer than

136   50 samples were excluded, resulting in 260 pollen taxa. The harmonisation and 'binning' of

137   sub-fossil and modern samples into 200-year long time windows resulted in a single dataset

138   of 4164 samples, of which 1610 are modern and 2554 are sub-fossil (from 105 sites). The

139   number of clusters was chosen through visual inspection of the dendrogram and

140   developments of measures that described intra-cluster compactness and inter-cluster

141   dissimilarity, on the basis of mean Euclidean distance scores.

142   A phytosociological approach was used to describe clusters identified within the pollen data.

143   The frequency of occurrence of taxa within each cluster was calculated, and the abundance

144   (pollen percentages) calculated using the median and inter-quartile range. Frequency of

145   occurrence was based on a five-point scale, where V represents a taxon occurring in 81-100%

146   of all samples in a cluster, IV = 61-80%, III = 41-60%, II = 21-40% and I = 0-20%.

147   Frequencies of V and IV indicate constant taxa, III common taxa, II occasional taxa and I

148   scarce taxa. This approach does not explicitly resolve issues of differential productivity of

149   pollen (Broström et al. 2008); the interpretation of the community data requires this to be

150   taken into consideration. Clusters have been given names on the basis of the constant taxa to

151   aid in description of the results. Detailed comparison between these names, modern

152   vegetation communities and pre-existing classification of pollen data (e.g. biomisation:

153   Prentice et al. 1996) is presented in Woodbridge et al. (in review). For each sample two

154   established Mediterranean 'human impact' indices have been calculated: OJC (Oleacaeae,

155   *Juglans*, *Castanea*: Mercuri et al., 2013a) and API (consisting of the human impact indicators

156   *Artemisia*, *Centaurea*, Cichorieae [syn. Lactuceae], *Plantago*, cereals, *Urtica*, *Trifolium*:

157   Mercuri et al., 2013b), and these indices have been summarised for each cluster.

158   In order to identify within-site pathways of vegetation change across consecutive time

159   windows the frequency of change from each pollen cluster to all other clusters was

160   calculated. These data have been visualised in chord diagrams, constructed in R using the

161   Circlize package (Gu 2014). Plots are circular in character, and use line width to indicate the

162   frequency by which samples change from one cluster grouping to another between adjacent

163   time windows, in similar fashion to that used in the "clock-face" diagrams of Walker (1970).

164   In order to aid visualisation, separate chord plots have been constructed for each cluster, and

165   are organised to indicate change routes towards a cluster (i.e. precursor clusters) and

166   destinations of change from clusters. Self-links (i.e. periods of stasis where a site does not

167   change cluster group across adjacent time windows) are ignored, as are gaps in individual site

168   records (i.e. where there are no dated samples within a time window).

169

170   **Results**

171   *Differentiation of clusters within the pollen dataset*

172   Sixteen groups of pollen samples have been extracted through the clustering analysis

173   following visual inspection of the dendrogram (Figure 2) and experimentation with

174   aggregation of samples at different levels. The dendrogram shows the aggregation of these

175   16 groups into eight meta-groups, and labelling of the clusters reflects both of these

176   groupings (i.e. meta-group 8 is an aggregation of four lower-level clusters, labelled 8.1, 8.2,

177   8.3 and 8.4). The phytosociological approach has been used to generate ecological meaning

178   for each grouping, drawing on the constant taxa that are present in each cluster. The results

179   of this are shown on Figure 3. Whilst importance is given to the range of values of each

180   constant taxon (shown as box-and-whisker plots with the interquartile range, median,

181   minimum and maximum values indicated), this is moderated by knowledge of relative pollen

182   productivity, including lower producers (typically the herbaceous taxa: Broström et al. 2008)

183   and high pollen producers such as Pinaceae (e.g. Stedingk et al. 2008).

184   Cluster 1 is an aggregation of four distinct groups that all appear to indicate open, disturbed

185   or human-modified vegetation as reflected in the OJC and API indices (Table 1; Figure 3).

186   Cluster 1.1 has as its constant taxa Poaceae, Ericaceae, Pinaceae, *Quercus* (evergreen taxa)

187   and Oleaceae, but is the least compact cluster that is produced according to the summary

188   Euclidean distance scores (Table 2). Cluster 1.2 has Oleaceae as a constant and dominant

189   taxon and has the highest OJC score (ranked 1/16), with a smaller number of co-constants

190   than other clusters, but including Poaceae and *Quercus* (evergreen taxa). Cluster 1.3 includes

191   as constants a series of steppe taxa, such as Chenopodiaceae, *Artemisia*, other open ground

192   taxa including Poaceae which results in the highest API score; Pinaceae and *Quercus*

193   (evergreen taxa) are also constant taxa. Cluster 1.4 has a suite of open ground taxa within the

194   constants, including Poaceae, *Plantago lanceolata*, Chenopodiaceae, Caryophyllaceae,

195   Asteraceae (subfamilies Asteroideae and Lactucoides), Ranunculaceae, and is one of only

196   two clusters to include Cerealia-type as a constant taxon. Cluster 1.4 is ranked 5/16 and 3/16

197   for the human impact indicator groups OJC and API.

198   Cluster 2 is a distinct group whether 16 or 8 clusters are derived from the dataset, and is both

199   very compact and distinct from other clusters as measured by average dissimilarity scores

200   (Table 2). The constant taxa (frequency class V) that dominate the assemblages in this group

201   are *Quercus* (evergreen taxa) and Poaceae, with Pinaceae, Oleaceae and Cistaceae notable in

202   frequency class IV. This group of samples is thus considered to represent evergreen oak

203   woodland/scrub. Cluster 3 includes within the constant taxa Cyperaceae (dominant),

204   Pinaceae, Poaceae and *Quercus* (deciduous taxa). This group is harder to interpret owing to

205   the possible ecological meanings of Cyperaceae, but Woodbridge et al. (in review) argue that

206   it represents a combination of lowland wetland and upland pasture. The cluster is ranked

207   6/16 on the API index on the basis of the median score for samples within the group.

208   Clusters 4 and 5 have as their dominant constant taxon Pinaceae (Figure 3), and all have

209   Poaceae and *Quercus* (deciduous taxa) as co-dominants. Overall these groupings have very

210   low OJC indices. The groupings are distinctive and separate out even at lower numbers of

211   clusters (Figure 2) and this is further reflected in the low intra-cluster average Euclidean

212   distance scores (Table 2). Cluster 4 has high values of Pinaceae, and most likely represents

213   pine forest. Cluster 5 is an aggregation of two groups. Cluster 5.1 has proportions

214   intermediate between clusters 4 and 5.2 (Figure 3), and cluster 5.2 has both a greater number

215   of constant taxa including open-ground indicators such as Chenopodiaceae, Caryophyllaceae

216   and Ranunculaceae, and higher proportions of open-ground indicators. Cluster 5.2 thus has a

217   higher API index, and is ranked 5/16. Cluster 5.1 is described as pine woods, and 5.2 as pine

218   steppe.

219   Cluster 6 has as its dominant constant *Quercus* (deciduous taxa), and is an aggregation of two

220   groups. Both groups have a large number of constant taxa (across both frequency classes V

221   and IV), and whilst the constant taxa in frequency class V are virtually identical, the groups

222   are differentiated on the proportions of these taxa, with 6.2 having lower proportions of

223   *Quercus* (deciduous taxa) and higher values across the open ground taxa. Cluster 6.2 is also

224   the only other group (other than 1.4) to include Cerealia-type as a constant, albeit at low

225   proportions, and is ranked 7/16 on the basis of the API index.

226   Cluster 7 is a distinct grouping (Figure 2), and is described as fir forest, as the constant

227   dominant taxon is *Abies*, and the other constant taxa are also largely woodland types. It has

228   the lowest rankings for both the OJC (15/16) and API (16/16) rankings. Cluster 8 is an

229   aggregation of four of the 16 clusters that are also largely dominated by woodland taxa, and

230   as a group have the lowest OJC and API scores. Cluster 8.3 appears to represent beech

231   woodland, 8.4 mesic forest, 8.1 alder woods and 8.2 non-Pinaceae coniferous forest. Whilst

232   these are compact and distinct clusters on the basis of the mean intra-cluster Euclidean

233   distance dissimilarity scores, the lower inter-cluster dissimilarity scores between the four sub-

234   clusters reflect their greater similarity (Table 2).

235   The spatial distribution of some common cluster groups for the early (11,000-8,200 BP), mid

236   (8,200-4,200 BP) and late (4,200 BP - present) Holocene are shown in Figure 4. The

237   Holocene subdivisions correspond to those defined by Walker *et al.* (2012). Some of the

238   vegetation clusters have strong regional distribution (Figure 4). Sclerophyll parkland (1.1),

239   for example, is most common in Iberia, and the central Mediterranean region shows the

240   clearest late Holocene increase in parkland/grassland (1.4). Although pine woods (5.1) are

241   widespread in southern France and Iberia in the early Holocene, this cluster group is absent in

242   the eastern Mediterranean at this time, and instead appears in the mid and late Holocene.

243   Deciduous oak woods (6.1) are most common in the eastern Mediterranean and

244   parkland/grassland (1.4) is common here throughout the Holocene, rather than showing a late

245   Holocene rise.

246   ***Transitions between clusters***

247   The change (transition) from one cluster group to another has been visualised in chord

248   diagrams (Figure 5). These show, in each case, the precursor groups (blue), and the

249   subsequent direction of change (green) for each cluster. Clusters that have few switches

250   either to or from them (<10) have been excluded from the plots (clusters 1.2, 2, 8.1, 8.2 and

251   8.3). Values on the plots are proportions of the total number of changes i.e. thicker lines

252   indicate more frequent changes in the direction indicated.

253   The key observation that can be made from the chord plots is that for the clusters that are

254   indicative of more open vegetation and those with higher OJC and API scores (clusters 1.1,

255   1.2, 1.3 and 1.4), there are no dominant transition routes from one cluster to another. These

256   clusters have a broad spectrum of pathways both to and from them across the range of sites

257   used here. As an example, transitions to cluster 1.1 (sclerophyllous parkland) come from all

258   other clusters with the exception of clusters 8.1 (alder woods) and 8.3 (beech woods). No

259   single route is dominant although some pathways are more common, including 1.2 (evergreen

260   shrubland) and 1.3 (sclerophyllous steppe/parkland). Onwards transitions for cluster 1.1 go

261   to one of nine (from 15) possible clusters. A similar pattern is observed for cluster 1.4

262   (parkland/grassland), with only cluster 5.1 (pine woods) not acting as a precursor and many

263   recorded 'destination' clusters. The same broad pattern is observed for cluster 3

264   (pasture/wetland). Almost all possible pathways to this cluster are recorded (with cluster 1.4

265   the most frequent, but not the dominant route).

266   In contrast, clusters that represent closed vegetation (i.e. forest, or natural as inferred from the

267   very low OJC and API scores) have fewer possible precursors. For example, cluster 4 (pine

268   forest) largely emerges from one of four clusters: 5.1 (pine woods), 5.2 (pine steppe), 8.4

269   (mesic forest) and 1.3 (sclerophyllous steppe/parkland). Cluster 6.1 (deciduous oak

270   woodland) comes from one of eight paths, but has dominant routes that include cluster 6.2

271   (deciduous oak parkland), 2 (evergreen broad-leaved shrubland) and 8.4 (mesic forest).

272   Cluster 7 (fir forest) only develops from cluster 8.4 (mesic forest) and cluster 8.3 (beech

273   forest). A greater number of possible destination clusters exist for these more closed

274   vegetation groups, with the exception of cluster 6.1 (deciduous oak woods) which largely

275   transitions to cluster 6.2 (deciduous oak parkland).

276

Discussion

Identifying natural and anthropogenic clusters

279   The value in using palaeoecology to inform resource and conservation managers in relation to

280   vegetation pathways and change under known past environmental change has been identified

281   by many authors (e.g. Birks 1996; Jackson and Blois 2014; Edwards et al. 2017). The

282   approach employed here has allowed us to gain novel insights into past groupings of pollen

283   samples, and by inference the nature of past vegetation, in a way that has not been attempted

284   previously. The results of the meta-analysis of pollen samples has identified clusters that fall

285   along a spectrum of human impact, from those assemblages that resemble more natural

286   communities, and those that show clear levels of anthropogenic transformation, supported by

287   the human impact indices OJC and API (Mercuri et al. 2013a,b). More natural vegetation

288   communities are those clusters characterised by deciduous (or sclerophyllous) woodland,

289   such as clusters 8.1-8.4 (alder, coniferous, beech and mesic forests), 6.1 (oak woods) and 7

290   (fir forest). The clusters in group 1 have the highest human impact scores from both

291   woodland cultivation (the OJC index) and pastoral and arable indicators (the API). These

292   communities are thus most likely to represent human-modified vegetation and thus

293   'anthromes' (cf Ellis 2011), as identified through novel groupings of pollen resulting from the

294   transformation of natural vegetation systems. Measures of ecosystem novelty are not

295   included here, although recent work has begun to demonstrate the timing and scale of

296   emergence of novel communities within Europe (Finsinger et al. 2017). A third broad

297   grouping is those clusters dominated by pine, although at this stage it is difficult to assess the

298   role of natural disturbance regimes (e.g. fire: Morales-Molino et al. 2013) in the development

299   or fragmentation of these pine groups.

300   Butzer (2005) describes two important land-use systems that emerge within the Holocene in

301   the Mediterranean: upland pastoral land use and Mediterranean polyculture. Mediterranean

302   polyculture represents a diverse set of land use management approaches in a single system,

303   including grain, animals and orchards (olive, walnut and grapes). According to Grove and

304   Rackham (2003) the scale of patches (fields, or blocks of individual types of land use) within

305   the polyculture system can be small and highly fragmented. Thus, any pollen samples from

306   within a cultural landscape may potentially include a variety of these different agricultural

307   land types, and still retain elements of less-disturbed vegetation. The use of key 'indicator'

308   taxa within the pollen clusters (the OJC and API indices) has allowed these transformed

309   vegetation communities to be identified. The use of two indicators of human impact further

310   supports the inferences: the Spearman's rho reveals a highly significant relationship between

311   the two indices ($rs[16] = 0.81$, $p < 0.000$).

312

313   *Transitions in Holocene Mediterranean vegetation*

314   The analysis of change between cluster groups can provide useful insights into the

315   development of the Mediterranean vegetation. Differences between natural and

316   anthropogenically-modified vegetation can be observed; for example, more natural clusters

317   (e.g. 4, 6.1, 7, 8.4: Figure 5) typically show only a small number of different precursor

318   groups. Thus pine forest (cluster 4: Figure 5) largely develops from pine woodland

319   (expansion of pine), pine steppe (development of woodland) and steppe parkland (early

320   Holocene expansion of pine). Deciduous oak woods (cluster 6.1: Figure 5) largely develops

321   from oak parkland (woodland expansion and development) and mesic forest (suggesting

322   competition with other species), although in a small number of cases it can replace

323   sclerophyllous vegetation types.

324   In contrast, those clusters identified as reflecting human-modified vegetation (within the

325   broad cluster groups 1 and 3) have a very wide set of potential pathways of change. The

326   transformation of natural vegetation towards these clusters (Figure 5), is thus not predictable.

327   This suggests that past societies did not preferentially alter one set of vegetation communities

328   over another (e.g. selective fragmentation of mesic forests over pine steppe), but appear to

329   have exploited and transformed a wide array of natural vegetation types. However,

330   vegetation changes are also influenced by the climatic limits and ecotones of different plant

331   communities. Previous comparisons of Mediterranean vegetation at 6000 BP and the

332   present-day confirm the transformation of much of the landscape (Collins et al. 2012). This

333   included not only an overall reduction in woodland cover (reflected in arboreal pollen

334   percentages) but also the existence, and transformation of, xeric communities between 6000

335   years ago and present. Collins et al. (2012) used this as evidence that the modern

336   Mediterranean is thus not principally a transformation from a largely mesic mid-Holocene

337   forest vegetation to the largely open and fragmented modern landscape. This is supported by

338   the findings presented here: anthropogenic clusters can emerge from many other vegetation

339   clusters.

340   A notable feature of the chord diagrams is reciprocity of pathways between clusters (Figure

341   5). For example, cluster 4 (pine forest) is frequently preceded by cluster 5.1 (pine woods).

342   Destinations from cluster 4 are also frequently cluster 5.1. A similar pattern can be observed

343   for clusters 6.1 and 6.2 (oak forest and oak parkland). This is in part a reflection of the

344   methodological basis of hierarchical clustering. All samples must fall within a discrete

345   cluster (as described below), and inevitably similar samples may fall within different classes

346   if they lie close to the divisions imposed on the data in the method. As vegetation rarely

347   changes abruptly in either space or time, temporal autocorrelation between samples in a site

348   may mean that adjacent assemblages are very similar, but 'flicker' (i.e. short-duration shifts)

349   between similar cluster groupings (e.g. oak parkland and oak forest, which are largely

350   differentiated on the basis of proportions of a similar group of constant taxa: Figure 3). Such

351   flickering has been observed in other classification methods with an arbitrary element (e.g.

352   pseudobiomisation, which uses an affinity score to assign pollen samples to predetermined

353   classes: Fyfe et al. 2010). Whilst this is the most likely explanation for the reciprocity

354   between groups, part of the pattern may also reflect changes in intensity of land use,

355   particularly when clusters move to, and from, clusters with higher or lower implied levels of

356   human modification such as oak parkland (API rank 7: Table 1) to parkland/grassland (API

357   rank 3: Table 1), or linked to periods of more/less frequent landscape burning (e.g. Kaniewski

358   et al. 2008). Butzer (2005) has described cycles of intensification and de-intensification

359   within the Mediterranean, which reflect expansion and contraction of complex societies

360 within a wider set of world systems. Within this framework of intensification/de-

361 intensification the expected pattern of Mediterranean vegetation would not be a monotonic

362 drive towards increased production and thus increased modification of vegetation. Phases

363 typified by re-expansion of more natural vegetation communities (perhaps on ~600 year time

364 scales, according to Butzer) would be expected. This may reflect movement towards, and

365 from, anthropogenic clusters (i.e. oak woodland to parkland with intensification, and back

366 with de-intensification). Butzer (2005) takes as an example the Lake Lerna pollen record

367 from the southern Argolid region of Greece (Jahns 1993). Cycles of disturbance (with

368 associated weed and cereals) and revertance (re-establishment of more natural vegetation)

369 can be clearly linked to the archaeological record from the region. The implications are that

370 whilst movement of samples between broadly similar clusters may reflect the characteristics

371 of the method, they may also reflect real variations in land use intensity.

372

The value of classification of pollen samples

374 The unconstrained organisation, or grouping, of data offers opportunities in the description of

375 past vegetation from pollen data, by allowing the inherent structure in the data to emerge,

376 particularly when a priori groupings in the data cannot be established (Felde et al. 2016).

377 Although previous work has also undertaken inductive data exploration such as that described

378 in this paper through classification (notably Huntley 1990) it remains an under-utilised

379 approach (Felde et al. 2014). The approach assumes that distinct groupings exist with the

380 dataset under study. Whilst there are in general clear and distinct vegetation communities

381 (and thus pollen assemblages) both in the present and the past Mediterranean landscape (e.g.

382 pine forests vs deciduous oak woodland) the gradual nature of vegetation change in both

383 space and time means that edges between groups may be 'blurred' (e.g. ecotones between

384 major vegetation communities). Further, pollen records integrate the signal of all vegetation

385 communities within the source area of the pollen. In highly heterogeneous landscapes this

386 may present particular challenges where pollen may be sourced from both highly modified

387 and more natural vegetation communities. In spite of these caveats, it has been possible

388 through this approach to recognise distinct Mediterranean vegetation types as reflected in sets

389 of distinctive pollen assemblages.

390   The results shown in the chord plots (Figure 5) are a first step in describing general patterns

391   of vegetation development across the Mediterranean basin, focusing on transitions from one

392   state to another. Future efforts might include comparison of Holocene pathways of change

393   with those from previous interglacials, which must reflect natural vegetation communities

394   and pathways of change (such as those described by Tzedakis 2007), to further elucidate

395   human impact on Holocene vegetation. Changes driven by extrinsic factors (e.g. shifts in

396   Mediterranean climate) might be expected to produce largely temporally consistent shifts in

397   vegetation, whilst intrinsic drivers (including site-specific human impacts) may produce

398   shifts that are more localised in both time and space. Clearly understanding such site-based

399   patterns demands an understanding of both regional differences in climate (e.g. Labuhn et al.

400   (in press) for the first millennium AD), and a detailed understanding of changing population

401   dynamics and land use systems from the archaeological and historical record (Perring and

402   Ellis 2013).

403   The unsupervised classification approach also has clear limitations, not least that all pollen

404   sites are treated equally. There has been no attempt to account for inter-site taphonomic

405   differences, such as scale of vegetation represented between small and large lakes, or

406   depositional biases between lakes and mires. Woodbridge et al. (2014) showed that whilst

407   such differences can produce noise in the analysis of large datasets, generally they do not

408   obscure the resulting signal that is obtained. Similarly, differential pollen production is not

409   accounted for (e.g. Fyfe et al. 2013; Marquer et al. 2014; Trondman et al. 2015). Placing

410   equal emphasis on frequency of occurrence of taxa and relative proportions of taxa, is an

411   attempt to alleviate this problem. Despite the limitations, there is still considerable value in a

412   data driven approach that is not restricted by the availability or applicability of PPEs and

413   allows major patterns in pollen datasets to be identified.

414

Conclusions

416   Ward's hierarchical clustering method has provided a simple approach to the division of a

417   large harmonised dataset from 105 sub-fossil pollen sequences and 1610 surface pollen

418   samples in the Mediterranean extracted from the EPD and EMPD. The division of the dataset

419   into sixteen groups has allowed the identification of clusters that are taken to represent more

420   natural vegetation communities (mesic, coniferous and sclerophyllous forest types), and more

421 open land cover types associated with human impact, supported by summary human impact

422 indicators. Some of these anthropogenic land cover types could be taken to represent

423 anthromes, that is, anthropogenically-modified biomes.

424 Clusters that are taken to represent more natural vegetation assemblages have limited

425 pathways of change between states. In contrast, those with a higher level of human

426 modification (as reflected in the OJC and API rankings) show numerous pathways to their

427 development. This implies widespread fragmentation of all community types as a

428 consequence of human impact, rather than selective transformation of particular types of

429 vegetation. Movement of sites from one cluster to a similar group and back may partly

430 represent an artefact of the method, but it is possible that some of these changes represent

431 cycles of intensification/de-intensification of land use, a phenomenon observed in the

432 archaeological record since prehistoric times.

433 Much more work is required to understand in detail the processes behind the transformation

434 and change in the Mediterranean pollen data represented here and this is a non-trivial task. It

435 will necessitate a comprehensive understanding of human land pressure around each site

436 including land use systems and population dynamics (drawn from archaeological and

437 historical sources), and the compilation of spatially-explicit climatic records. The ability to

438 describe vegetation at the sub-continental scale using approaches such as that presented here

439 is clear, and a next step is to develop the independent (i.e. non-pollen) datasets that can help

440 understand the patterns that emerge.

References

Birks HJB (1996) Contributions of Quaternary palaeoecology to nature conservation. Journal of Vegetation Science 7:89-98

Birks HJB, Lotter AF, Juggins S, Smol JP (2012) Tracking Environmental Change Using Lake Sediments, Volume 5: Data Handling and Numerical Techniques. Springer, Dordrecht

Broström A, Nielsen AB, Gaillard M-J, Hjelle K, Mazier F, Binney H, Bunting M-J, Fyfe RM, Meltsov V, Poska A, Räsänen S, Soepboer W, Stedingk H, Suutari H. Sugita, S (2008) Pollen productivity estimates – the key to landscape reconstructions. Vegetation History and Archaeobotany 17:461-478

Butzer KW (2005) Environmental history in the Mediterranean world: cross-disciplinary investigation of cause-and-effect for degradation and soil erosion. Journal of Archaeological Science 32:1773-1800

Collins PM, Davis BAS, Kaplan JO (2012) The mid-Holocene vegetation of the Mediterranean region and southern Europe, and comparison with the present day. Journal of Biogeography 39:1848-1861

Davis BAS, Zanon M, Collins P, Mauri A, Bakker J, Barboni D, Barthelmes A, Beaudouin C, Birks, HJB, Bjune AE, Bozilova E, Bradshaw RHW, Brayshay BA, Brewer S, Brugiapaglia E, Bunting J, Conner SE, de Beaulieu J-L, Edwards KJ, Ejarque A, Fall P, Florenzano A, Fyfe R, Galop D, Giardini M, Giesecke T, Grant MJ, Guiot J, Jahns S, Jankovská V, Juggins S, Karmann M, Karpinska-Kolaczek M, Kolaczek P, Kuhl N, Kunes P, Lapteva EG, Leroy SAG, Leydet M, Lopez Saez JA, Masi A, Matthias I, Mazier F, Meltsov V, Mercuri AM, Miras Y, Mitchell FJG, Morris JL, Naughton F, Nielsen AB, Novenko E, Odgaard B, Ortu E, Overballe-Petersen MV, Pardoe HS, Peglar SM, Pidek I, Sadori L, Seppa H, Severova E, Shaw H, Swieta-Musznicka J, Theuerkauf M, Tonkov S, Veski S, van der Knaap WO, van Leeuwen J, Woodbridge J, Zimny M, Kaplan JO (2013) The European Modern Pollen Database (EMPD) project. Vegetation History and Archaeobotany 22:521-530

Edwards KJ, Fyfe RM, Jackson ST (2017) The first 100 years of pollen analysis. *Nature Plants* 3. doi:10.1038/nplants.2017.1

471 Ellis EC (2011) Anthropogenic transformation of the terrestrial biosphere. Philosophical

472   Transactions of the Royal Society of London A: Mathematical, Physical and

473   Engineering Sciences 369:1010-1035

474 Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N (2010) Anthropogenic

475   transformation of the biomes, 1700 to 2000. Global Ecology and Biogeography 19:589-

476   606

477 Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the

478   world. Frontiers in Ecology and the Environment 6:439–447

479 Felde VA, Bjune AE, Grytnes JA, Birks HJB (2014) A comparison of novel and traditional

480   numerical methods for the analysis of modern pollen assemblages from major

481   vegetation–landform types. Review of Palaeobotany and Palynology 210:22-36

482 Felde VA, Hooghiemstra H, Torres-Torres V, Birks JBH (2016) Detecting patterns of change

483   in a long pollen-stratigraphic sequence from Funza, Colombia – a comparison of new

484   and traditional numerical approaches. Review of Palaeobotany and Palynology 234:94-

485   109

486 Finsinger W, Giesecke T, Brewer S, Leydet M (2017) Emergence patterns of novelty in

487   European vegetation assemblages over the past 15 000 years. Ecology Letters 20:336-

488   346

489 Fyfe RM, de Beaulieu J-L, Binney H, Bradshaw RHW, Brewer S, Le Flao A, Finsinger W,

490   Gaillard M-J, Giesecke T, Gil-Romera G, Grimm EC, Huntley B, Kunes P, Kuhl N,

491   Leydet M, Lotter AF, Tarasov PE and Tonkov S (2009) The European Pollen Database:

492   past efforts and current activities. Vegetation History and Archaeobotany 18:417-424

493 Fyfe RM, Roberts CN, Woodbridge J (2010) A pollen-based pseudo-biomisation approach to

494   anthropogenic land cover change. The Holocene 20:1165-1171

495 Fyfe RM, Twiddle C, Sugita S, Gaillard M-J, Barratt P, Caseldine CJ, Dodson J, Edwards KJ,

496   Farrell M, Froyd C, Grant MJ, Huckerby E, Innes JB, Shaw H, Waller M (2013) The

497   Holocene vegetation cover of Britain and Ireland: overcoming problems of scale and

498   discerning patterns of openness. Quaternary Science Reviews 73:132-148

499 Fyfe RM, Woodbridge J, Roberts CN (2015) From forest to farmland: pollen-inferred land

500   cover change across Europe using the pseudobiomization approach. Global Change

501   Biology 21:1197-1212

Gaillard MJ, Birks HJB, Emanuelsson U, Karlsson S, Lagerås P, Olausson D (1994) Application of modern pollen/land-use relationships to the interpretation of pollen diagrams—reconstructions of land-use history in south Sweden, 3000-0 BP. Review of Palaeobotany and Palynology 82:47-73

Gaillard M-J, Sugita S, Mazier F, Trondman, A-K, Broström A, Hickler T, Kaplan JO, Kjellström E, Kuneš P, Lemmen C, Olofsson J, Smith B, Strandberg G, Fyfe R, Nielsen AB, Alenius T, Balakauskas L, Barnekow L, Birks HJB, Bjune A, Björkman L, Giesecke T, Hjelle K, Kalnina L, Kangur M, van der Knaap WO, Koff T, Lagerås P, Latalowa M, Leydet M, Lechterbeck J, Lindbladh M, Odgaard B, Peglar S, Segerström U, von Stedingk H, Seppä H (2010) Holocene land cover reconstructions for studies on land cover-climate feedbacks. Climate of the Past 6:483-499

García-Madrid AS, Molina JA, Cantó P (2014) Classification of habitats highlights priorities for conservation policies: The case of Spanish Mediterranean tall humid herb grasslands. Journal for Nature Conservation 22:142-156.

Giesecke T, Davis B, Brewer B, Finsinger W, Wolters S, Blaauw M, de Beaulieu J-L, Fyfe RM, Gaillard M-J, Gil-Romera G, van der Knaap WO, Kuneš P, Kühl N, van Leeuwen JFN, Leydet M, Lotter AF, Semmler M, Bradshaw RHW (2014) Towards mapping the late Quaternary vegetation change of Europe. Vegetation History and Archaeobotany 23:75-86

Grove AT, Rackham, O (2003) The nature of Mediterranean Europe: an ecological history. Yale University Press, Yale

Gu Z (2014) Circlize implements and enhances circular visualisation in R. Bioinformatics 30:2811-2812

Hijmans RJ, Cameron SE, Parra JL, Jones PJ, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology **25**:1965-1978

Huntley B (1990) European vegetation history: Palaeovegetation maps from pollen data-13000 yr BP to present. Journal of Quaternary Science 5:103-122.

Izdebski A, Holmgren K, Weiberg E, Stocker SR, Büntgen U, Florenzano A, Gogou A, Leroy AAG, Luterbacher J, Martrat B, Masi A, Mercuri AM, Montagna P, Sadori L, Schneider A, Sicre MA, Triantaphyllou M, Xoplaki E (2015) Realising consilience:

533   How better communication between archaeologists, historians and natural scientists can

534   transform the study of past climate change in the Mediterranean. Quaternary Science

535   Reviews 136:5–22.

536   Jackson ST, Blois JL (2015) Community ecology in a changing environment: Perspectives

537   from the Quaternary. Proceedings of the National Academy of Sciences 112:4915-4921

538   Jahns S (1993) On the Holocene vegetation of the Argive Plain (Peloponnese, southern

539   Greece). Vegetation History and Archaeobotany 2:187-203.

540   Kaniewski D, Paulissen E, De Laet V, Waelkens M (2008) Late Holocene fire impact and

541   post-fire regeneration from the Bereket basin, Taurus Mountains, southwest Turkey.

542   Quaternary Research 70:228-239.

543   Labuhn I, Finné M., Izdebski A, Roberts N, Woodbridge J (in press) Climatic changes and

544   their impacts in the Mediterranean during the first millennium CE. *Late Antique*

545   *Archaeology*

546   Leydet M (2007-2017) The European Pollen Database.

547   http://www.europeanpollendatabase.net/. Accessed 26 May 2016

548   Marquer L, Gaillard M-J, Sugita S, Trondman A-K, Mazier F, Nielsen AB, Fyfe RM,

549   Odgaard B, Alenius T, Birks HJB, Bjune AE, Christiansen J, Dodson J, Edwards KJ,

550   Giesecke T, Herzschuh U, Kangur M, Lorenz S, Poska A, Schult M, Seppä H (2014)

551   Holocene changes in vegetation composition in northern Europe: why pollen-based

552   quantitative reconstructions matter? Quaternary Science Reviews 90:199-216

553   Mercuri AM, Bandini Mazzanti M, Florenzano A, Montecchi MC, Rattighieri E (2013a)

554   *Olea, Juglans* and *Castanea*: the OJC group as pollen evidence of the development of

555   human-induced environments in the Italian peninsula. Quaternary International 303:24-

556   42

557   Mercuri AM, Mazzanti M, Florenzano A, Montecchi MC, Rattighieri E, Torri P (2013)

558   Anthropogenic pollen indicators (API) from archaeological sites as local evidence of

559   human-induced environments in the Italian peninsula. Annali Di Botanica 3:143-153

560   Morales-Molino C, García-Antón M, Postigo-Mijarra JM, Morla C (2013) Holocene

561   vegetation, fire and climate interactions on the westernmost fringe of the Mediterranean

562   Basin. Quaternary Science Reviews 59:5-17

563 Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which
564   algorithms implement Ward's criterion? Journal of Classification 31:274-295

565 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
566   RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2016) vegan:
567   Community Ecology Package. R package version 2.4-2. https://CRAN.R-
568   project.org/package=vegan

569 Ozendaa P, Borela JL (2000) An ecological map of Europe: why and how? Ecology
570   323:983–994

571 Perring MP, Ellis EC (2013) The extent of novel ecosystems: long in time and broad in space.
572   In: Hobbs J, Higgs ES, Hall CM (eds) Novel ecosystems: Intervening in the new
573   ecological world order, Wiley-Blackwell, Chichester, pp 66-80

574 Prentice C, Guiot J, Huntley B, Jolly D, Cheddadi R (1996) Reconstructing biomes from
575   palaeoecological data: a general method and its application to European pollen data at 0
576   and 6 ka. Climate Dynamics 12:185-194

577 Prentice IC, Jolly D, BIOME 6000 Participants (2000) Mid-Holocene and glacial-maximum
578   vegetation geography of the northern continents and Africa. Journal of Biogeography
579   27:507–519

580 Radeloff VC, Williams JW, Bateman BL, Burke KD, Carter SK, Childress ES, Cromwell KJ,
581   Gratton C, Hasley AO, Kraemer BM (2015) The rise of novelty in ecosystems.
582   Ecological Applications 25:2051-2068

583 Roberts N (2013) The Holocene: an environmental history. John Wiley and Sons, London

584 Roberts N, Stevenson T, Davis B, Cheddadi R, Brewster S, Rosen A (2004) Holocene
585   climate, environment and cultural change in the circum-Mediterranean region. In:
586   Batterby RW, Gasse F, Stickley CE (eds) Past climate variability through Europe and
587   Africa. Springer, Netherlands, pp 343-362

588 Ruddiman W, Fuller D, Kaplan J, Ellis E, Kutzbach J, Vavrus S, Tzedakis P, Roberts N, Fyfe
589   R, He F, Lemmen C, Woodbridge J (2016) Late Holocene climate: natural or
590   anthropogenic? Review of Geophysics 54:93-118

591 Stedingk H, Fyfe RM, Allard A (2008) Pollen productivity estimates for the reconstruction of
592   past vegetation at the forest-tundra ecotone. The Holocene 18:323-332

593 Sugita S (2007a) Theory of quantitative reconstruction of vegetation I: pollen from large sites
594 REVEALS regional vegetation composition. *The Holocene* 17:229-241

595 Sugita S (2007b) Theory of quantitative reconstruction of vegetation II: all you need is
596 LOVE. The Holocene 17:243-257

597 Trondman A-K, Gaillard M-J, Sugita S, Mazier F, Fyfe R, Lechterbeck J, Marquer L, Nielsen
598 AB, Twiddle C, Barratt P, Birks HJB, Bjune AE, Caseldine C, David R, Dodson J,
599 Dörfler, W, Fischer E, Giesecke T, Hultberg T, Kangur M, Kuneš P, Latalowa M,
600 Leydet M, Lindbaldh M, Mitchell F, Odgaard B, Peglar SM, Persson T, Rösch M, van
601 der Knaap P, van Geel B, Smith A, Wick L (2015) Pollen-based land-cover
602 reconstructions for the study of past vegetation-climate interactions in NW Europe at
603 0.2 k, 0.5 k, 3 k and 6 k years before present. Global Change Biology 21:676-697

604 Tzedakis PC (2007) Seven ambiguities in the Mediterranean palaeoenvironmental narrative.
605 Quaternary Science Reviews 26:2042-2066

606 von Post L (1946) The prospect for pollen analysis in the study of the earth's climatic history.
607 New Phytologist 45:193-193

608 Vannière B, Colombaroli D, Roberts N (2010) A fire paradox in ecosystems around the
609 Mediterranean. PAGES News 18:63-65.

610 Walker D (1970) Direction and rate of change of some British post-glacial hydroseres. In:
611 Walker D, West RG (eds) Studies in the vegetation history of the British Isles.
612 Cambridge University Press, Cambridge, pp. 117-39.

613 Walker MJC, Berkelhammer M, Björck S, Cwynar LC, Fisher DA, Long AJ (2012) Formal
614 subdivision of the Holocene Series/Epoch: a Discussion Paper by a Working Group of
615 INTIMATE (Integration of ice-core, marine and terrestrial records) and the
616 Subcommission on Quaternary Stratigraphy (International Commission on
617 Stratigraphy). Journal of Quaternary Science 27:649-659

618 Ward JH (1963) Hierarchical grouping to optimize an objective function. Journal of the
619 American Statistical Association 58:236-244

620 Williams JW, Blois JL, Shuman BN (2011) Extrinsic and intrinsic forcing of abrupt
621 ecological change: case studies from the late Quaternary. Journal of Ecology 99:664-
622 677

623 Williams JW, Jackson ST (2007) Novel climates, no analog communities, and ecological

624   surprises. Frontiers in Ecology and the Environment 5:475-482

625 Williams M, Zalasiewicz J, Haff PK, Schwägerl C, Barnosky AD, Ellis EC (2015) The

626   Anthropocene biosphere. The Anthropocene Review 2:196-219

627 Woodbridge J, Fyfe RM, Roberts N (2014) A comparison of remotely-sensed and pollen-

628   based approaches to mapping Europe's land cover. Journal of Biogeography 41:2080-

629   2092

630 Woodbridge J, Roberts N, Fyfe RM (in review) Perpetual change? Holocene vegetation and

631   land cover dynamics in the Mediterranean from pollen data. Journal of Biogeography

632

633 Figure 1: Location of sites from the Mediterranean region from the European Pollen Database
634 and European Surface Sample Pollen Database within geographical (A) and climatological
635 space (B: EPD sites; C: EMPD surface samples). In (B) and (C) overall Mediterranean
636 climate is described using a kernel density plot of all grid cells within the Mediterranean
637 region (data taken from Hijmans et al. 2005).

638
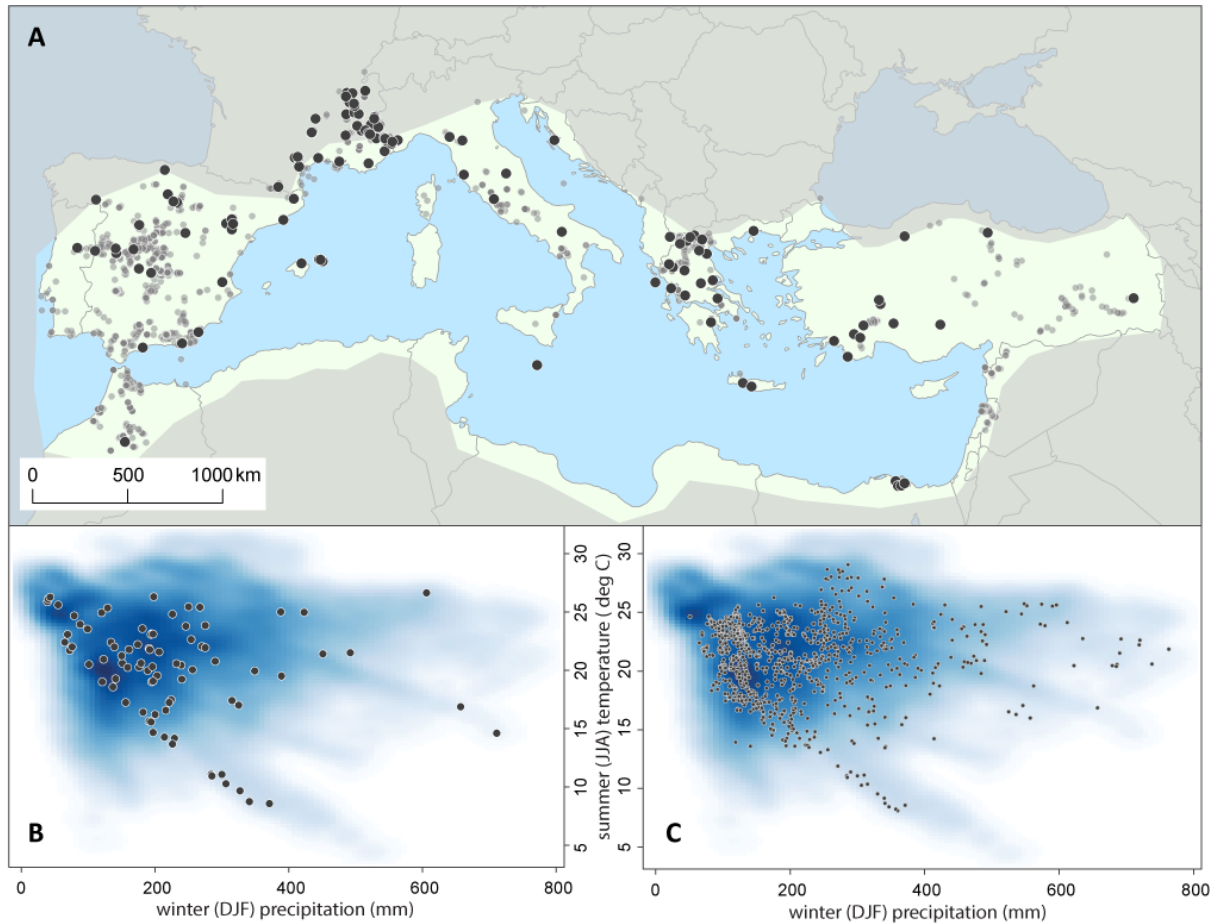
639

640

641   Figure 2: Dendrogram derived from hierarchical clustering of the pollen samples using
642   Ward's method. The final 16 derived clusters are indicated with solid boxes; the higher level
643   of aggregation to 8 'meta-clusters' is indicated with dotted boxes.
644
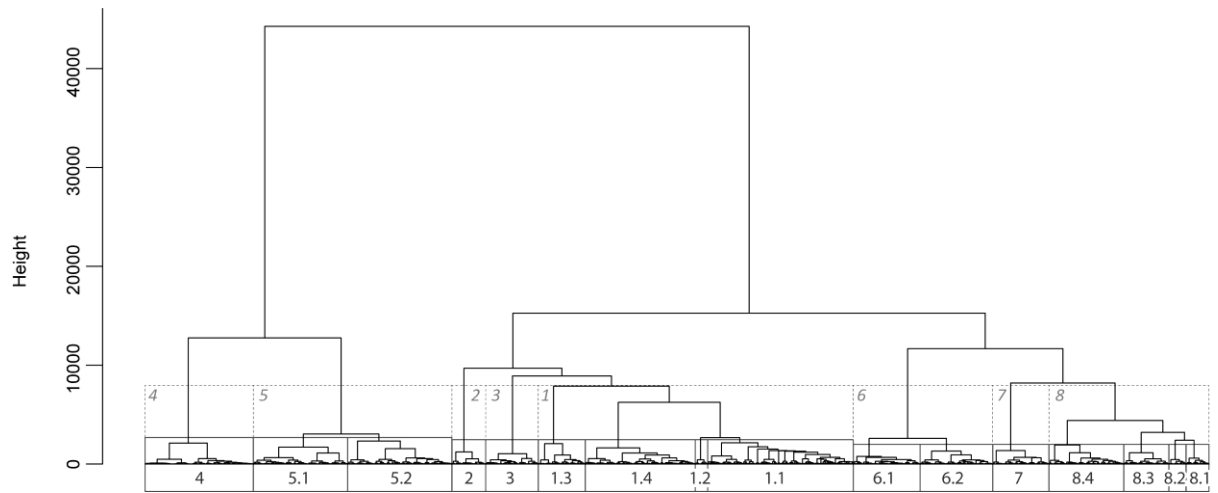
645

646

647 Figure 3: Constant taxa in each of the sixteen cluster groupings derived from Ward's
648 hierarchical clustering method. Only constant taxa (frequency groups V (dark shading) and
649 IV (mid-grey shading)) are shown. Boxes indicate the interquartile range of each taxon
650 within the cluster, with the median, maximum and minimum values also shown.
651
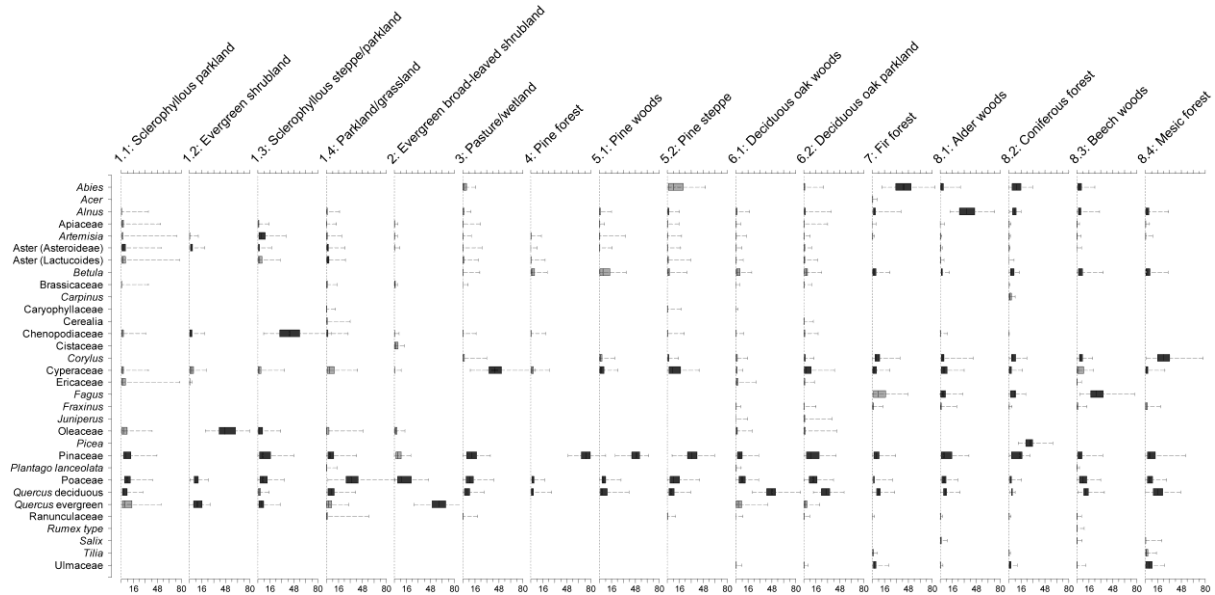
652
653

654 Figure 4a,b: Spatial and temporal distribution of cluster groups 1.1, 1.4, 5.1 and 6.1. Samples
655 are aggregated into broad time windows to show patterning between the early-, mid- and late-
656 Holocene.
657
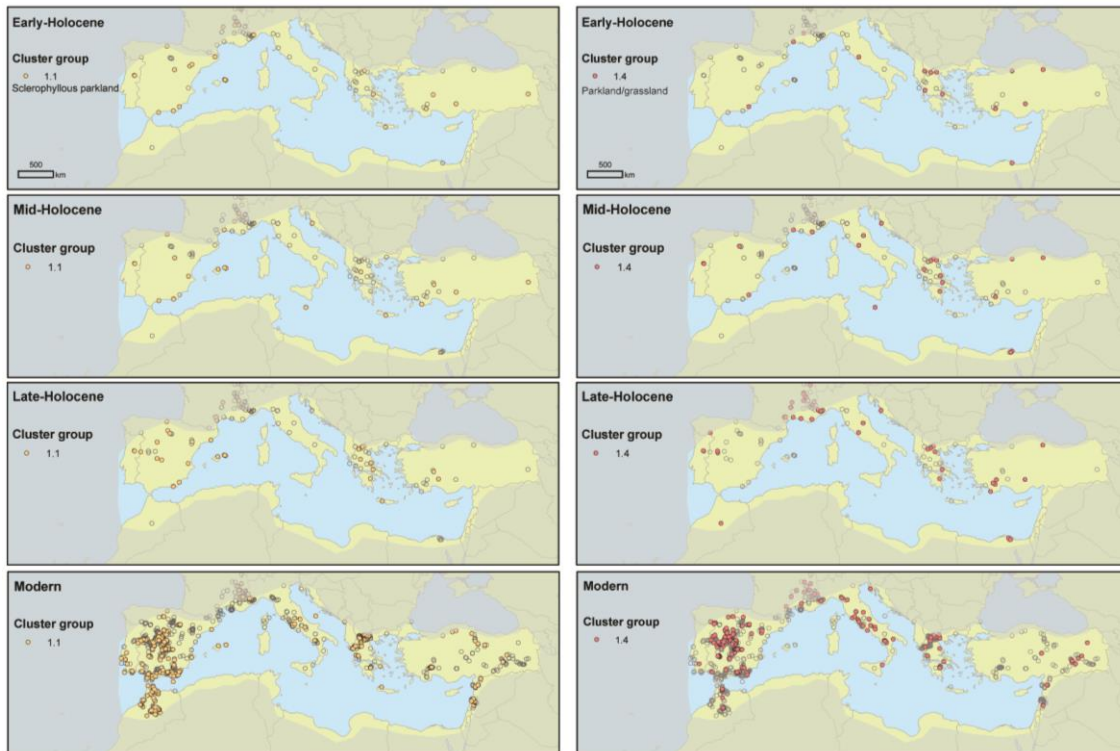
658
659
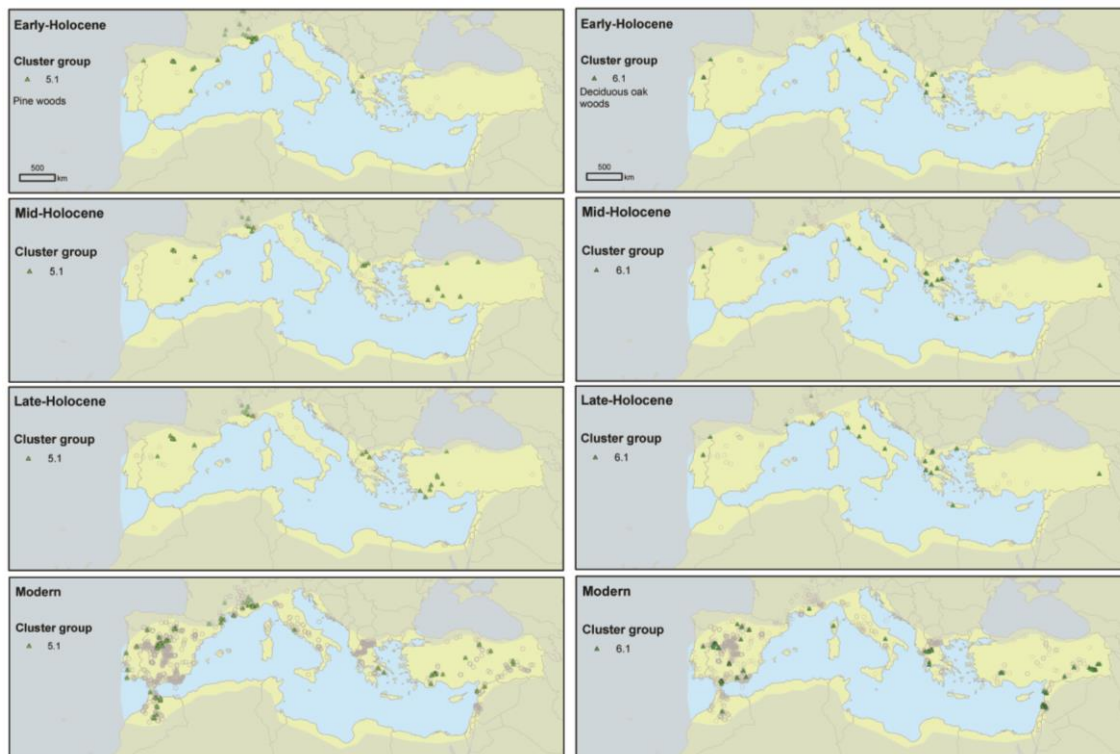
660
661

662 Figure 5: Chord diagrams for each pollen cluster indicating frequency and direction of
663 change for each group. Blue lines flow towards the cluster, indicating the precursor groups.
664 Green lines flow from each cluster, and indicate the groupings to which samples in each site
665 switch. Values of 't' and 'f' on each plot indicate total number of switches to and from each
666 cluster; 'nc' indicates percentage of samples that do not switch across consecutive time
667 windows.
668

1.1: sclerophyllous parkland
1.3: sclerophyllous steppe
1.4: parkland/grassland
  3: pasture/wetland
  4: pine forest
5.1: pine woods
5.2: pine steppe
6.1: deciduous oak woods
6.2: deciduous oak parkland
  7: fir forest
8.4: mesic forest
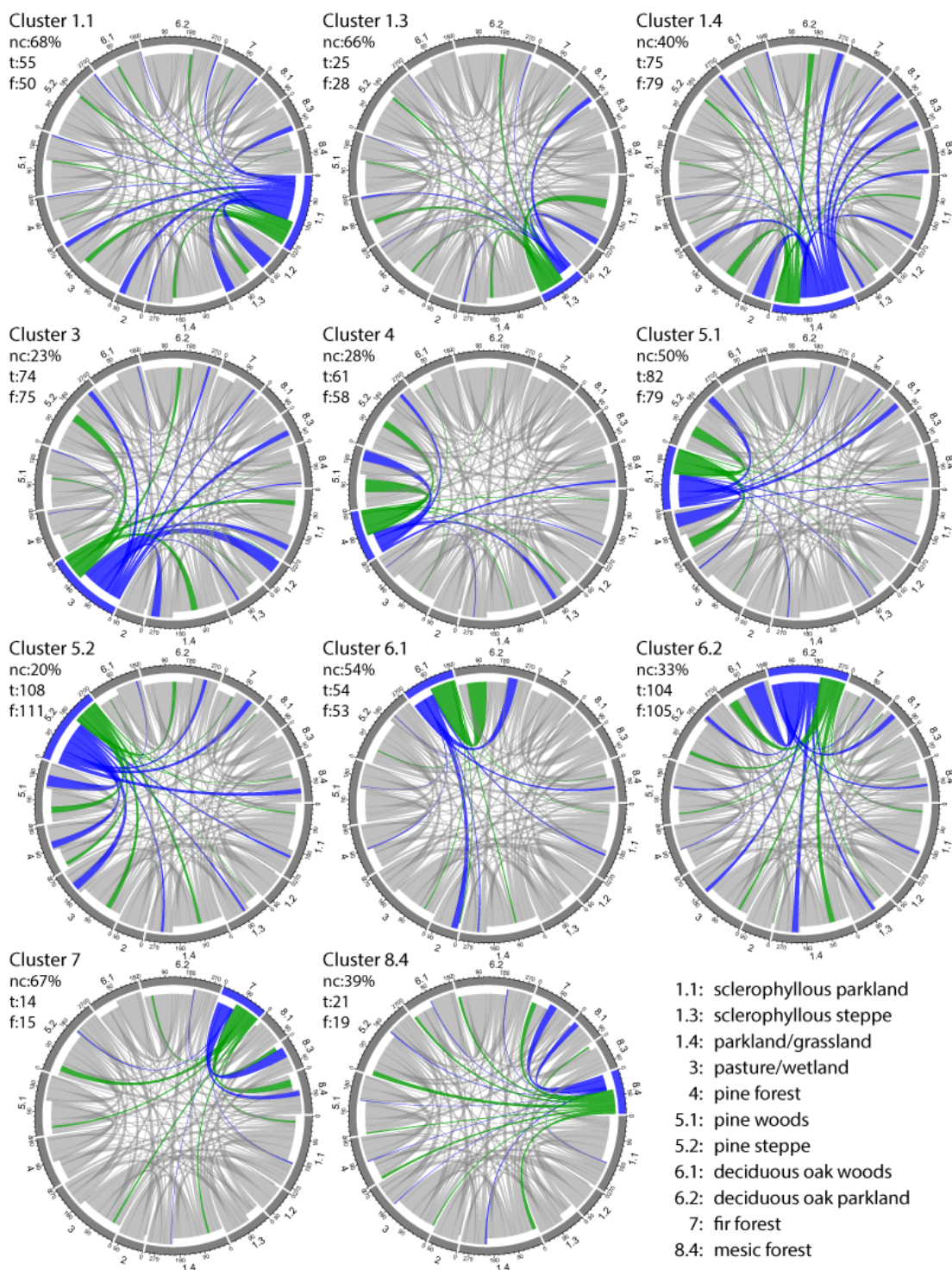
669
670

671 Table 1: Descriptions of the cluster groups, including dominant taxa and summaries of the human
672 impact indices used (OJC: Oleaceae, *Juglans*, *Castanea sensu* Mercuri et al., 2013a; API: *Artemisia*,
673 *Centaurea*, Lactuceae [syn. Cichorieae], *Plantago*, cereals, *Urtica*, *Trifolium sensu* Mercuri et al.,
674 2013b) and ranking of clusters by human impact indices.

Cluster	Cluster name (dominant taxa)	%OJC index median (IQR)	OJC rank	%API index median (IQR)	API rank
1.1	Sclerophyllous parkland (Poaceae-Quercus-Oleaceae-herbaceous and sclerophyll taxa)	2.42 (0.56-8.47)	2	7.89 (2.45-22.57)	2
1.2	Evergreen shrubland (Oleaceae-Quercus evergreen-Poaceae-herbaceous and sclerophyllous taxa)	46.78 (39.05-61.8)	1	6.27 (3.03-9.76)	4
1.3	Sclerophyllous steppe/parkland (Chenopodiaceae-Poaceae-Artemisia-Quercus evergreen-Oleaceae)	1.85 (0.45-5.88)	4	8.64 (4.96-16.3)	1
1.4	Parkland/grassland (Poaceae, Quercus deciduous, Asteraceae)	0.97 (0.23-3.44)	5	6.32 (3.4-12.29)	3
2	Evergreen broad-leaved shrubland (Quercus evergreen, Poaceae)	2.14 (0.79-4.41)	3	2.42 (1.44-4.05)	8
3	Pasture/wetland (Cyperaceae, Poaceae, Quercus deciduous)	0.42 (0-2.46)	8	3.26 (0.96-6.96)	6
4	Pine forest (Pinaceae)	0.03 (0-0.73)	13	1.57 (0.79-3.08)	13
5.1	Pine woods (Pinaceae-Poaceae-Quercus deciduous)	0.14 (0-0.68)	9.5	2.15 (0.94-4.56)	9
5.2	Pine steppe (Pinaceae-Poaceae-Quercus deciduous-Artemisia)	0.13 (0-0.96)	11	3.44 (1.53-7.52)	5
6.1	Deciduous oak woods (Quercus-Poaceae-Pinaceae)	0.64 (0.07-2.74)	6	1.62 (0.85-3.07)	12
6.2	Deciduous oak parkland (Quercus-Poaceae-Pinaceae-sclerophyll taxa)	0.43 (6.06-2.07)	7	3.04 (13.13-6.16)	7
7	Fir forest (Abies-Quercus deciduous-Corylus-Pinaceae)	0 (0-0)	15	0.5 (0.17-1.06)	16
8.1	Alder woods (Alnus-Quercus deciduous-Cyperaceae)	0.06 (0-0.55)	12	1.91 (0.79-3.55)	10
8.2	Coniferous forest (Picea-Pinaceae-Abies)	0 (0-0.3)	15	1.2 (0.74-2.04)	14
8.3	Beech woods (Fagus-Quercus deciduous-Poaceae-Corylus)	0.14 (0-0.67)	9.5	1.72 (0.96-3.24)	11
8.4	Mesic forest (Corylus-Quercus deciduous-Poaceae-Pinaceae-Ulmaceae)	0 (0-0)	15	1.05 (0.45-2.86)	15

675 Table 2: Summary of dissimilarity matrix scores for each cluster, and dissimilarity scores showing
676 difference between clusters. Values are based on average Euclidean distance between all samples
677 within, or between, clusters. Higher scores indicate a greater degree of dissimilarity.

	1.1	1.2	1.3	1.4	2	3	4	5.1	5.2	6.1	6.2	7	8.1	8.2	8.3	8.4
1.1	**53.4**															
1.2	62.6	**28.0**														
1.3	63.4	70.8	**38.9**													
1.4	52.7	62.1	62.3	**34.1**												
2	67.4	72.9	78.5	66.5	**26.7**											
3	61.1	70.6	69.7	55.0	78.3	**30.5**										
4	78.2	91.8	84.3	80.0	94.0	79.6	**18.6**									
5.1	60.0	74.7	68.5	60.2	78.0	62.1	35.3	**28.2**								
5.2	53.8	67.8	62.7	50.2	73.3	51.3	51.0	36.9	**34.1**							
6.1	61.2	72.1	72.1	57.3	76.3	66.1	85.6	65.4	59.7	**25.7**						
6.2	52.0	64.3	63.1	45.6	70.4	53.2	72.5	52.3	45.9	35.6	**31.4**					
7	64.2	74.3	73.6	63.5	80.7	66.0	84.6	66.5	55.0	65.0	58.0	**30.8**				
8.1	59.5	70.5	69.3	57.5	76.3	60.7	79.0	61.3	54.5	62.1	52.6	61.5	**32.5**			
8.2	55.2	67.1	65.3	54.8	74.0	59.1	72.7	54.5	47.9	60.3	50.4	51.6	51.2	**23.9**		
8.3	54.8	66.6	65.7	50.7	73.4	57.6	79.0	58.8	51.1	53.1	45.2	53.0	51.7	46.0	**29.9**	
8.4	55.9	68.5	66.9	53.1	75.0	59.4	76.5	57.0	51.4	49.9	44.1	56.9	53.6	49.0	46.3	**35.0**

678

679