

Epidemic Diffusion Network of Spain: a mobility model to characterize transmission routes of disease. A COVID-19 case study.

Javier Del Águila Mejía (✉ javidelaguila@gmail.com)

Universidad Autonoma de Madrid Facultad de Medicina <https://orcid.org/0000-0001-5697-5980>

David García-García

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Ayelén Rojas-Benedicto

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Nicolás Rosillo

Hospital Universitario 12 de Octubre

María Guerrero-Vadillo

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Marina Peñuelas

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Rebeca Ramis

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Diana Gómez-Barroso

Centro Nacional de Epidemiología, Instituto de Salud Carlos III

Juan de Mata Donado-Campos

Universidad Autonoma de Madrid Facultad de Medicina

Research Article

Keywords: Mobility, Transmission, Networks, Infectious Diseases, Epidemiology, Spatial epidemiology, disease modelling

Posted Date: June 27th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1740822/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **TITLE:** Epidemic Diffusion Network of Spain: a mobility model to characterize
2 transmission routes of disease. A COVID-19 case study.

3
4 **AUTHORS:**

5
6 Javier Del-Águila-Mejía^{1,2,3}

7 David García-García^{2,4}

8 Ayelén Rojas-Benedicto^{2,4}

9 Nicolás Rosillo⁵

10 María Guerrero-Vadillo²

11 Marina Peñuelas²

12 Rebeca Ramis^{2,4}

13 Diana Gómez-Barroso^{2,4}

14 Juan de Mata Donado-Campos^{1,4,6}

15 (1) Departamento de Medicina Preventiva y Salud Pública y Microbiología, Facultad de
16 Medicina. Universidad Autónoma de Madrid. C. Arzobispo Morcillo, 4, 28029
17 Madrid

18 (2) Centro Nacional de Epidemiología, Instituto de Salud Carlos III. Calle de Melchor
19 Fernández Almagro, 5, 28029. Madrid, Spain.

20 (3) Servicio de Medicina Preventiva, Hospital Universitario de Móstoles. Calle Río
21 Júcar, s/n, 28935. Móstoles, Spain

22 (4) Consorcio de Investigación Biomédica en Red de Epidemiología y Salud Pública
23 (CIBERESP). Instituto de Salud Carlos III. Calle Monforte de Lemos 3-5. 28029.
24 Madrid, Spain.

25 (5) Servicio de Medicina Preventiva, Hospital Universitario 12 de Octubre. Avenida de
26 Córdoba, s/n, 28041. Madrid, Spain.

27 (6) Instituto de Investigación Sanitaria del Hospital Universitario La Paz (IdiPAZ)
28 Universidad Autónoma de Madrid, Madrid, Spain.

29 Corresponding author: Javier Del Águila Mejía. javidelaguila@gmail.com

30 Co-senior authors: Juan de Mata Donado Campos and Diana Gómez-Barroso

31
32 **ORCID:**

33 Javier Del-Águila-Mejía: <https://orcid.org/0000-0001-5697-5980>

34 Ayelén Rojas-Benedicto: <https://orcid.org/0000-0003-3197-0731>

35 Nicolás Rosillo: <https://orcid.org/0000-0002-4046-8092>

36 María Guerrero-Vadillo: <https://orcid.org/0000-0003-0575-789X>

37 Marina Peñuelas: <https://orcid.org/0000-0002-5909-7957>

38 Rebeca Ramis: <https://orcid.org/0000-0001-6154-9142>

39 Diana Gómez-Barroso. <https://orcid.org/0000-0001-7388-1767>

41 **ABSTRACT**

42 **Background:** Human mobility drives geographical diffusion of airborne infectious
43 diseases at different scales, and the COVID-19 pandemic has made mobility data
44 widely available thanks to technological advances. Mobility has been used to
45 investigate COVID-19 but few studies focus on mobility itself. We used public data
46 from February 14th 2020 to May 9th 2021 in Spain to characterize mobility patterns
47 between the 52 provinces to build the Epidemic Diffusion Network using network
48 science methods to study geographical diffusion phenomena.

49 **Results:** A total of 135 (out of 2.264) connections between the 52 provinces were
50 identified as the most relevant and a weighted, directed network was built. Madrid,
51 Valladolid and Araba/Álaba scored the highest degree and strength. Shortest routes
52 (most likely path between two points) were obtained between all provinces. A total of
53 7 mobility communities were found with a network modularity of 63%. COVID-19
54 cumulative incidence in 14 days (CI14) reflected community structure, with provinces
55 of the same cluster behaving more similar when compared with the outside ones.
56 Shortest routes of diffusion between all 52 provinces were obtained.

57 **Conclusions:** Mobility patterns in Spain are governed by a small number of high-
58 flow connections that remain constant in time and seem unaffected by seasonality
59 or restrictions. Most of the travels happen within communities, that does not
60 completely represent political borders, thus indicating the importance of coordination
61 between administrations when addressing health emergencies. The most likely
62 geographical spread pattern in Spain is wave-like with occasional long-distance
63 jumps, like a small-world network. This information could be of use to enhance
64 preparedness and response plans with interventions targeted to relevant locations
65 before the disease spreads, and to inform better models of disease diffusion at the
66 country level.

67

68 **KEYWORDS:**

69 Mobility, Transmission, Networks, Infectious Diseases, Epidemiology, Spatial
70 epidemiology, disease modelling

71

72 BACKGROUND

73 The COVID-19 disease took only a few months since its emergence in the Chinese province
74 of Wuhan to become a global pandemic affecting all the World regions (1). As other respiratory
75 viruses, its main route of infection is direct person-to-person contact between an infectious host
76 and a susceptible individual, and therefore COVID-19 dynamics are shaped mainly by the way both
77 individuals and populations interact with each other (2). This implies a multi-scale nature of
78 transmission dynamics in which human mobility plays a key role explaining local outbreaks and
79 epidemics (3), geographical diffusion (4, 5) and international spread of pandemics (6, 7).

80 We have seen a surge in mobility information availability mainly due to the innovative development
81 of big data techniques (8–11). Studies of the impact of mobility on COVID-19 dynamics were soon
82 available in China (12–14), Italy (15, 16), Spain (17) or USA (5, 9, 18) among others (19). Mobility
83 data has been used to model geographical diffusion (12–15, 17–19), epidemic growth (5, 13, 16) or
84 to evaluate non pharmaceutical interventions in place (5, 12).

85 Most of these studies treat mobility as another explaining variable within a multiplicity of different
86 methodologies. But few of them focus on how the subjacent mobility matrix impacts on the
87 epidemic diffusion (14, 20).

88 Network science has long studied the diffusion of epidemics in contact or travel networks (20).
89 Predicting risks posed by transmission events between a small number of regions (cities, provinces,
90 states) can be relatively straightforward. These are usually proportional to the number of people
91 who move between points and the characteristics of the transmitted phenomena (21). But when it
92 comes to the study of epidemic diffusion in an entire network consisting of several locations
93 interconnected, a new complexity arises and a multiplicity of possible routes, distances and
94 geographical patterns emerge (14, 7).

95 The first COVID-19 cases were detected in Spain in March 2020 (22). As many other countries, a
96 national-wide lockdown was issued by the government to successfully contain the first wave. During
97 2020 and 2021, several waves took place (22), with mobility restrictions as one of the main control
98 measures in place.

99 The aim of the present study is to characterize mobility patterns in Spain, defining a Mobility Matrix
100 (MM) native to Spain. This MM is then used to investigate COVID-19 dynamics during 2020 and 2021
101 and to explore whether mobility patterns are linked to epidemiological outcomes.

102 METHODS

103 Data sources

104 Mobility data is publicly provided by the Ministry of Transportation, Mobility
105 and Urban Agenda (MITMA) (23). We defined Spain's 52 provinces as the area of study
106 (Map 1). Data was aggregated at that level in the form of daily unique travels between
107 provinces, from February 14th 2020 to May 9th 2021.

108 COVID-19 data for Spain is available from the Instituto de Salud Carlos III public
109 information dashboard (22). This information is gathered through the National
110 Surveillance Network (RENAVE, in Spanish). Demographic information is provided by
111 the Instituto Nacional de Estadística (INE) (24).

112 Map 1. Political map of Spain

113 Mobility Matrix of Spain and Shortest Distances maps

114 A description of mobility patterns in Spain is first carried out to characterize
115 behavioral, seasonal, and geographical properties. Let F_{nm} be the flux of daily total
116 travelers emanating from a province m to any different province n . Supplementary
117 materials detail the process and the rationale to turn these fluxes into the fraction P_{nm}
118 of people leaving m and going to n among all people leaving m (7) (Eq. 1)

$$119 \quad P_{nm} = \frac{F_{nm}}{\sum_{n \neq m} F_{nm}}.$$

120 P_{nm} fractions represent the distribution of the relative importance of each destination
121 province n connected with m . We define an adjacency matrix containing all P_{nm} pairs
122 for the 52 provinces and name it Mobility Matrix (MM). As stated in Supplementary
123 Materials, this MM contains a qualitative and quantitative key representation of the
124 mobility patterns in Spain, regardless of variation in time, season, or measures in place.
125 To further study transmission phenomena within this MM, all percentages P_{nm} are
126 turned into a distance-like unit: Effective Distance (ED), as proposed and tested by (7):

$$127 \quad d_{nm} = 1 - \log(P_{nm}).$$

128 ED is a measure of proximity: the larger P_{nm} , the lesser d_{nm} and therefore the closer
129 is node n from m in terms of disease diffusion (7). Complex connection patterns
130 emerge from the matrix P_{nm} and its d_{nm} derivate (Supplementary material), since
131 EDs allow to mathematically compute paths along the MM that imply multi-step routes
132 due to the aggregative property of logarithms compared with the conditional
133 probability of percentages (7). We can find the shortest distances between two
134 provinces D_{nm} as the minimal sum of d_{nm} required to reach node n from m .

135 For example, if $d_{BA} = 6$ but $d_{CA} = 2$ and $d_{BC} = 2$, any travelling disease would
136 arrive from A to B faster by transitioning through C ($D_{BA} = 4$) rather than directly from
137 A to B. The set of shortest paths between pairs of provinces within the MM are obtained
138 using the classical Dijkstra's algorithm (25) implemented in the *igraph* R package. For
139 each province, a map of all the shortest incoming and outgoing D_{nm} is created.

140 Epidemic Diffusion Network and measures

141 From the MM we define a network (26) to represent and uncover these complex
142 mobility patterns that might drive diffusion events. To model the network, only

143 relevant connections (e.g., those with P_{mm} over 10%, see Supplementary material) are
144 selected. When applied to the study of disease diffusion dynamics, we name this
145 network the “Epidemic Diffusion Network” (EDN).

146 In a network, the degree of a node is the number of edges connected to it (incoming and
147 outgoing), a representation of connectiveness. In weighted networks, the strength of a
148 node is the sum of the weights of each edge connected to the node i.e., the weighted
149 degree. Degree and strength are local measures that account only for neighboring
150 properties of nodes. Betweenness, on the other hand, sums up the number of times the
151 node lies within the path between any other two nodes, thus scoring nodes that act as
152 bridge or intersecting points (27, 28).

153 Network substructure analysis was performed to detect communities (29). A found
154 community must be understood as a mobility cluster, in epidemiological terms. This
155 implies that nodes are connected in such way that movement (flow) within the
156 community is much more likely. The percentage of total movements happening within
157 communities rather than between them is known as the modularity score. Clustering of
158 nodes happens hierarchically and will be represented with a dendrogram.

159 Network structure and COVID-19 dynamics

160 Since mobility drives the geographical diffusion of COVID-19, it would be
161 expected that provinces within the same community, which share the most of their
162 mobility fluxes, also share more similar disease dynamics. These dynamics are summed
163 up in the cumulative incidence in 14 days (CI14) epidemic curves and therefore we
164 compare their trajectories, shapes and magnitude along the study period.

165 We adopt an analytical approximation by proposing a measure of curve proximity. We
166 compute the absolute difference in daily CI14 between any two provinces m, n for a
167 given day t

168

$$d_{mn,t} = |CI14_{m,t} - CI14_{n,t}|,$$

169 where $CI14_{m,t}$ denotes the CI14 at province m on day t . The mean $\overline{d_{mn}}$ of these daily
170 differences over the time under study is the global proximity measure between the
171 disease dynamics of any two provinces. The lower $\overline{d_{mn}}$ value, the more similar
172 COVID-19 dynamics two provinces have experienced.

173 The $52*52$ possible $\overline{d_{mn}}$ are calculated between each province and 1) the rest of the
174 provinces within its community and 2) the rest of the provinces outside. All obtained
175 means are compared with a T test with the null hypothesis of no differences in $\overline{d_{mn}}$,
176 i.e., being part of a community does not imply a difference in disease dynamic for the
177 studied period.

178 All analysis have been made using R version 4.0.3 (30) and the network package *igraph*
179 (31).

180 RESULTS

181 Mobility Matrix of Spain and Shortest Distance maps

182 Daily aggregated total mobility from February 14th, 2020, to May 9th, 2021, is
183 shown in figure 1. The first week (Feb 14th to Feb 21st 2020) is considered the
184 reference mobility value before pandemic began. A first sharp decline in March 2020
185 happens after the first declaration of the stay-at-home order, lifted in June when
186 mobility almost reached pre-pandemic values. By the end of August, it declines again
187 not to recover to the end of our study period, even after the second state of alarm came
188 into force from October 24th to May 9th.

189 **Figure 1. Daily total travels between provinces and 14-day cumulative**
190 **incidence in Spain, 14th Feb 2020 – 9th May 2021**

191 *Cumulative incidence in 14 days per 100.000 inhabitants.*

192 **Epidemic Diffusion Network**

193 The 52 provinces are the network's nodes and, between them, 135 weighted edges (26)
194 represent the main MM connections as defined in Supplementary material. An example
195 of shortest distances maps is shown in figure 2. The modelled network is presented in
196 Fig 3. Madrid scores the highest degree (11), followed by Valladolid (9), Araba/Álava
197 (8), Málaga (8), Sevilla (8) and Valencia (8). Madrid, Valladolid and Araba/Álava also
198 record the highest strength (25.7, 25.3 and 20.2). Network flow betweenness, on the
199 other hand, concentrates in provinces of the southernmost community. Mean (SD)
200 values for the entire network are 5.2 (2) degree, 12.5 (5.5) strength, 318.1 (408.9) flow
201 betweenness (Table S1).

202 **Figure 2. Shortest distances maps from (outgoing) and to (incoming) two**
203 **provinces, Huesca and Madrid**

204 *A) Shortest distances to Madrid; B) Shortest distances from Huesca*

205 A total of 7 communities are detected within the network (Fig 3). The biggest
206 community is the southernmost consisting in a total of 12 provinces connected to the
207 outside only by two provinces. On the other hand, the smallest communities are the two
208 composed by Canary Islands (two provinces) and Galicia (4 provinces). The first is
209 dependent only from Madrid as main connection. The other four communities
210 represent different geographical regions. Highest scoring provinces in terms of
211 degree/strength are well divided among communities. Network's modularity is 63%.
212 Hierarchical clustering dendrogram and qualitative conclusions derived from it are
213 depicted in Supplementary materials.

214 **Figure 3. Effective distance network (EDN) representation of network measures**
215 **and map of communities**

216 *A) vertex size is proportional to incoming/outgoing connections (degree). Each edge represents*
217 *a connection. Communities are represented by color. B) Map of Spain with provinces colored by*
218 *community belonging.*

219 **Network structure and COVID-19 dynamics**

220 COVID-19 CI14 epidemic curves by provinces grouped by community can be
221 seen in figure 4. Provinces in the same network community seem to share more similar
222 COVID-19 incidence, represented by their curves' trajectories, shapes and magnitude.
223 This is confirmed by the fact that all provinces have a statistically significant lower
224 mean daily difference $\overline{d_{ij}}$ with other provinces within their same community
225 compared with outer ones. Excluding Canarias, with only two provinces, the higher
226 difference score is for the provinces in the southern community (106.3 vs 68.4; $p <$
227 $.000$). On the contrary, the smaller difference is found with the central-eastern
228 community encompassing Madrid and Valencia (116 vs 97.2; $p < .000$) (Figure 5).

229 **Figure 4. CI14 per 100.000 inhabitants by province, grouped by community**
230 **during the study period**

231 **Figure 5. Mean daily differences between CI14 curves within communities' vs**
232 **outside community for every detected community**

233 *Note: all mean differences between groups for all communities were statistically significant*

234 **DISCUSSION**

235 We used raw mobility data gathered during the COVID-19 pandemic to curate
236 and develop a Mobility Matrix (MM) in the form of an adjacency matrix between the 52

237 provinces of Spain that was used to build a network model to unveil complex properties
238 that drive geographical diffusion phenomena linked to mobility (such as, but not limited
239 to, epidemics). This network has been described and linked to COVID-19 dynamics
240 between March 2020 and May 2021.

241 Mobility has been analyzed and described in successive layers of complexity:
242 first, by depicting the raw fluxes of daily travelers. Secondly, by turning them into a
243 distance-like variable and unveiling the constant relationships and patterns emerging
244 from the data. And third, by building up a network model to further understand
245 diffusion phenomena within that network. The latter was mainly achieved through
246 hierarchical clustering for community detection and the shortest distances maps. Each
247 of these layers of study allowed us to gain a share of insight into the complexity of
248 human mobility in one specific way.

249 In February 2020, before the COVID-19 crisis began, more than 5 million people
250 travelled daily between any two provinces in Spain. With the first stay-at-home order
251 in March, when just essential work-related travels were allowed, only 600.000 people
252 still travelled. By the summer restrictions had been fully lifted and mobility came back
253 to normal only to start declining again as the second COVID-19 wave was happening, to
254 never recover in the studied period.

255 The MM reveals that diffusion phenomena are governed by major fluxes rather than
256 small variations i.e., holidays, and how measures put in place were able to reduce the
257 number of people travelling but not their usual destinations. Bilateral relationship and
258 major fluxes between provinces are constant. Each province has a unique set of main
259 neighbours and a way of interacting with them. These constitute the properties of
260 provinces which will determine their role in diffusion phenomena as the nodes of a
261 network, connected by their effective distances on the edges.

262 Finally, the network itself reveals some emergent properties such as modularity, the
263 presence of communities of mobility and the estimation of all the shortest paths

264 between provinces. It is in this layer where we will mainly discuss epidemiological
265 implications of our findings for disease transmission.

266 Disease epidemiology in networks is a growing field with a solid background, especially in
267 animal ecology (32, 33). Network measures have been studied in this context, centrality measures
268 being the most relevant for infectious diseases study (32). They unveil the role and position each
269 node holds within the network, thus translating it into epidemiological risks, e.g., of infectivity,
270 exposure, or disease facilitators (32, 27, 20). These properties are directly given by the MM and it is
271 of no surprise that the higher scoring provinces in terms of degree and strength are capital regions
272 (Madrid, Valladolid, Araba/Álava, Sevilla and Valencia) in their respective Autonomous
273 Communities. In fig 2 a hub-based configuration of mobility can be observed for most communities
274 in which one or two nodes act as main attractor and “receiver”.

275 The network shows a configuration of hubs that connect with smaller provinces rather than
276 between them, acting more like local gravitating poles for adjacent lesser provinces rather than
277 drivers of the whole network, in accordance with the high modularity score observed (63%).

278 However, the province of Barcelona, 2nd largest city of Spain, is an exception. Contrary to the rest
279 of communities, the northeastern shows a completely different behavior in which no main hub
280 emerges. Degree and strength are more equally distributed, and it is Zaragoza (another capital city)
281 which scores higher, but not too different from the rest. We are unable to explain these findings,
282 but the resulting sub-network points towards a differentiated diffusion dynamic compared with the
283 rest of Spain.

284 The seven communities found are consistent with the administrative division of Spain (figure 3), yet
285 there is no perfect fit. This is especially true for the 2nd and 3rd largest central communities, with
286 Madrid acting as main bridge between two mobility clusters and several Autonomous Communities
287 (including the Canary Islands). Hierarchical clustering (Fig 4) shows how the southern community
288 conforms an independent area from the rest of Spain. Whereas, the remaining areas are more
289 closely related, making hard to contain the spread of a disease, especially when involving the capital,
290 Madrid. However, at the same time, the north-east community shares a different relationship
291 between its components, with no clear hub emerging. This region has been specially affected during
292 successive COVID-19 waves, for no known reason so far. These findings could shed light in this
293 matter in future studies.

294 We have demonstrated how communities impacted COVID-19 dynamics at the province level as
295 CI14 evolution showed less differences between provinces sharing a mobility community for all
296 communities independently. This effectively links mobility patterns and spatial diffusion of COVID-
297 19 in Spain. Together, our findings point towards a locally wise approach to measures like mobility
298 restrictions.

299 Our network, the EDN, is directed, weighted and thresholded but many different networks
300 can be modeled for disease transmission depending on underlying assumptions, giving rise to what
301 is called the network's structure (27, 34). The network in which a disease spreads highly determines
302 the evolution and dynamics it will experience (20). It links the individual behaviour to broad
303 population effects through either contact patterns during an outbreak (35) or the spatial
304 (geographical) component, as well as distances, paths or clustering degrees (20). Among many
305 idealized types of networks, small-world networks (36) are characterized by a high clustering and
306 short paths lengths, such as the one we find in the EDN, but also by the presence of random links
307 that connect otherwise distant elements of the network.

308 This implies that diffusion would take the form of a spreading wave from an origin source, with the
309 ability to reach distant elements quickly (20). These wave-like patterns were also observed by
310 Brockman and Helbing in the international air-travel network during the pandemics of H1N1 in 2009
311 or SARS in 2003 (7). A recent study of COVID-19 during the summer 2020 in Spain points towards
312 the same conclusion: a rapid spread in a region (north-east of Spain) but then a sudden jump to
313 Madrid from where it reached the rest of Spain (37). This could be associated with the effective
314 distance map emanating from the first provinces to experience the epidemic in July 2020 (Fig 2).

315 High modularity was also found in our network, meaning that the disease is less likely to quickly
316 spread through the network (38), yet other authors consider this effect rather humble, especially in
317 highly contagious diseases which would only be affected with a modularity over 80%, which is rare
318 (39). In spite of this, modularity is still a determining factor to consider for control strategies in
319 epidemiology since in the presence of a less transmissible disease, chances are that appropriate
320 measures increase the so-called structural trapping of the network i.e., the ability to prevent the
321 diffusion of the disease from one community to another (39).

322 Reducing human mobility has been the main tool for COVID-19 control worldwide. At the
323 geographical level, both diffusion patterns and local epidemics are affected by the way we

324 collectively move. This network approach has provided valuable insight into the properties of
325 mobility in Spain and their implications in epidemiological terms regarding disease diffusion.

326 By knowing the main travel routes and patterns we can move from the simple metric of mobility
327 reduction, which is only useful when affecting all the population, a situation every day less likely to
328 happen. Potentially susceptible locations could implement measures in advance under the event
329 of local outbreaks taking place. Three levels of relation have been presented: direct relation
330 (effective distance), indirect relation (shortest paths) and community structure. It is by combining
331 information from all of them that preventive measures should be tackled.

332 The relation between the network position and disease dynamics has been pointed out but needs
333 further research. We lack evidence or a solid framework to decide which measures are the most
334 useful (32). A particular measure can be calculated in different networks yet not every measure
335 will mean the same in all network structures (27).

336 All the relations between network components and disease epidemiology need to be further tested
337 with more advanced models including SIR-like models within the network to fully characterize
338 relations and implications of these finding. Simple SEIR models are able to fully capture the
339 complexity of these dynamics when paired with the adequate subjacent network (3). The massive
340 amount of epidemiological and social information made available during the COVID-19 pandemic
341 poses a unique opportunity to address these gaps in knowledge and help us advance towards a
342 more generalized theoretical framework if the field.

343 Limitations

344 This is an ecological study using only aggregated mobility without any individual information which
345 may determine the implications for disease transmission events. Mobility data was gathered
346 through mobile phone information, so not all travels might have been captured. The administrative
347 unit used in this work, the province, might hide subtle patterns at smaller level e.g., municipalities,
348 that could lead to different conclusion. The reasons for choosing the province were pragmatic:
349 public health units usually operate at this level and restrictions are mostly imposed by provinces
350 rather than municipalities. Regarding COVID-19 incidence, limitations are those of surveillance data:
351 reporting bias, different interpretation by date, testing and reporting capacities changing over time
352 and the aggregation of information.

353 Network science when applied to epidemiology is an emerging field. More research is needed to
354 effectively extract definitive conclusions. Network measures are affected by its design and methos
355 implemented, different groups may reach different conclusion and no general framework for
356 making such decisions exist.

357 **Conclusion and recommendations**

358 We have unveiled stable patterns of human mobility that allowed us to model geographical
359 transmission routes between provinces in Spain and to study mobility properties through network
360 methods. This information can be adopted by public health authorities to inform future measures
361 to tackle outbreaks or health emergencies of infectious nature. Routes of transmission allow to
362 preventively act in case of necessity, and properties of the provinces (nodes) allow to design specific
363 interventions adapted to each region's characteristics. Infectious diseases do not know about
364 borders, and mobility communities encompass multiple administrations in Spain, thus requiring a
365 strong coordination between regional and central authorities. Including mobility information as
366 routine part of epidemiological surveillance can yield important benefits and requires further study.

367

368 **DECLARATIONS**

369

370 **Ethics approval and consent to participate**

371 Not applicable

372

373 **Consent for publication**

374 Not applicable

375

376 **Availability of data and materials**

377 The datasets used and/or analysed during the current study are publicly available from the
378 following sources:

- 379 • Instituto de Salud Carlos III. Situación y evolución de la pandemia de COVID-19
380 en España. Available from: URL: <https://cnecovid.isciii.es/covid19/>
- 381 • Ministerio de Transportes, Movilidad y Agenda Urbana. Estudio de movilidad con
382 Big Data. Available from: URL: <https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data>

383

384

385 **Competing interests**

386 The authors declare that they have no competing interests

387

388 **Funding**

389 This research has been financed by Carlos III Health Institute (ISCIII) under the project
390 COV20-00881

391

392 **Authors' contributions**

393 JDAM is the main author and took part in all phases of the project, from designing, data
394 processing and analysis, writing and editing. DGG took part in analysis design and
395 manuscript editing. ARB and NRR are responsible for graphics and code development. MP
396 and MG V are part of the COVID-19 data team. RM provided consultation during all phases
397 of the project. DGB and JMDC are the senior authors.

398

399 **Acknowledgements**

400 The authors wish to thank all the colleagues who provided valuable insight and support
401 during the project at its many phases.

402

403

404

405

406

407

408 BIBLIOGRAPHY

409

- 410 1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Geneva; 2020.
411 Available from: URL: <https://covid19.who.int> [cited 2021 Sep 22].
- 412 2. Brauer F, van den Driessche P, Wu J. *Mathematical Epidemiology*. Springer: Springer Berlin
413 Heidelberg; 2008. (vol 1945).
- 414 3. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D et al. Mobility network models
415 of COVID-19 explain inequities and inform reopening. *Nature* 2020;1–8. Available from: URL:
416 <https://www.nature.com/articles/s41586-020-2923-3>.
- 417 4. Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L et al. Human mobility and the
418 spatial transmission of influenza in the United States. *PLoS Comput Biol* 2017;
419 13(2):e1005382.
- 420 5. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility
421 patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet*
422 *Infectious Diseases* 2020; 20(11):1247–54. Available from: URL:
423 [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30553-3/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30553-3/fulltext).
- 424 6. Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, Vespignani A. Modeling the spatial
425 spread of infectious diseases: the GLObal Epidemic and Mobility computational model. *J*
426 *Comput Sci* 2010; 1(3):132–45.
- 427 7. Brockmann D, Helbing D. The hidden geometry of complex, network-driven contagion
428 phenomena. *Science* 2013; 342(6164):1337–42.
- 429 8. Grantz KH, Meredith HR, Cummings DAT, Metcalf CJE, Grenfell BT, Giles JR et al. The use of
430 mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat Commun*
431 2020; 11(1):1–8. Available from: URL: [https://www.nature.com/articles/s41467-020-18190-](https://www.nature.com/articles/s41467-020-18190-5)
432 5.
- 433 9. Kang Y, Gao S, Liang Y, Li M, Rao J, Kruse J. Multiscale dynamic human mobility flow dataset
434 in the U.S. during the COVID-19 epidemic. *Sci Data* 2020; 7(1):1–13. Available from: URL:
435 <https://www.nature.com/articles/s41597-020-00734-5>.

- 436 10. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, Nadai M de et al. Mobile phone data
437 for informing public health actions across the COVID-19 pandemic life cycle. *Sci. Adv.* 2020;
438 6(23):eabc0764.
- 439 11. Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C et al. COVID-19 outbreak
440 response, a dataset to assess mobility changes in Italy following national lockdown. *Sci Data*
441 2020; 7(1).
- 442 12. Jiang J, Luo L. Influence of population mobility on the novel coronavirus disease (COVID-
443 19) epidemic: based on panel data from Hubei, China. *glob health res policy* 2020; 5(1):1-10.
- 444 13. Kraemer MUG, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM et al. The effect of human
445 mobility and control measures on the COVID-19 epidemic in China. *Science* 2020;
446 368(6490):493-7.
- 447 14. Gibbs H, Liu Y, Pearson CAB, Jarvis CI, Grundy C, Quilty BJ et al. Changing travel patterns in
448 China during the early stages of the COVID-19 pandemic. *Nat Commun* 2020; 11(1):5012.
- 449 15. Cartenì A, Di Francesco L, Martino M. How mobility habits influenced the spread of the
450 COVID-19 pandemic: Results from the Italian case study. *Sci Total Environ* 2020; 741:140489.
- 451 16. Cintia P, Fadda D, Giannotti F, Pappalardo L, Rossetti G, Pedreschi D et al. The relationship
452 between human mobility and viral transmissibility during the COVID-19 epidemics in Italy;
453 2020 Jun 4. Available from: URL: <https://arxiv.org/pdf/2006.03141>.
- 454 17. Mazzoli M, Mateo D, Hernando A, Meloni S, Ramasco JJ. Effects of mobility and multi-
455 seeding on the propagation of the COVID-19 in Spain; 2020.
- 456 18. Kissler SM, Kishore N, Prabhu M, Goffman D, Beilin Y, Landau R et al. Reductions in
457 commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New
458 York City. *Nat Commun* 2020; 11(1):4674. Available from: URL:
459 <https://www.nature.com/articles/s41467-020-18271-5>.
- 460 19. Kishore N, Kahn R, Martinez PP, Salazar PM de, Mahmud AS, Buckee CO. Lockdowns result
461 in changes in human mobility which may impact the epidemiologic dynamics of SARS-CoV-2.
462 *Sci Rep* 2021; 11(1):6995. Available from: URL: [https://www.nature.com/articles/s41598-](https://www.nature.com/articles/s41598-021-86297-w)
463 021-86297-w.
- 464 20. Keeling MJ, Eames KTD. Networks and epidemic models. *J R Soc Interface* 2005; 2(4):295-
465 307.

- 466 21. Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA. Population flow drives spatio-temporal
467 distribution of COVID-19 in China. *Nature* 2020; 582(7812):389–94.
- 468 22. Instituto de Salud Carlos III. Situación y evolución de la pandemia de COVID-19 en España.
469 Available from: URL: <https://cnecovid.isciii.es/covid19/>.
- 470 23. Ministerio de Transportes, Movilidad y Agenda Urbana. Estudio de movilidad con Big Data.
471 Available from: URL: [https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-](https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data)
472 [big-data](https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data) [cited 2021 Apr 15].
- 473 24. Instituto Nacional de Estadística (INE). INEbase. Available from: URL:
474 <https://www.ine.es/dyngs/INEbase/listaoperaciones.htm>.
- 475 25. Dijkstra EW. A note on two problems in connexion with graphs. *Numerische Mathematik*
476 1959 [cited 2021 Jun 25]; 1(1):269–71.
- 477 26. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. The architecture of complex
478 weighted networks. *Proc Natl Acad Sci USA* 2004 [cited 2021 Apr 17]; 101(11):3747–52.
479 Available from: URL: <https://www.pnas.org/content/101/11/3747>.
- 480 27. Bell DC, Atkinson JS, Carlson JW. Centrality measures for disease transmission networks.
481 *Social Networks* 1999; 21(1):1–21. Available from: URL:
482 <https://www.sciencedirect.com/science/article/pii/S0378873398000100>.
- 483 28. Takemoto K, Oosawa C. Introduction to Complex Networks: Measures, Statistical
484 Properties, and Models. In: Dehmer M, Basak SC, editors. *Statistical and machine learning*
485 *approaches for network analysis*. Hoboken, N.J.: Wiley; 2012. p. 45–75.
- 486 29. Pons P, Latapy M. Computing communities in large networks using random walks (long
487 version); 2005 Dec 12. Available from: URL: <https://arxiv.org/pdf/physics/0512106>.
- 488 30. R Core Team. R: A language and environment for statistical computing. R Foundation for
489 Statistical Computing. Vienna, Austria. Available from: URL: <https://www.R-project.org/>.
- 490 31. Gabor Csardi, Tamas Nepusz. The igraph software package for complex network research.
491 *InterJournal* 2006; *Complex Systems*:1695. Available from: URL: <https://igraph.org>.
- 492 32. Silk MJ, Croft DP, Delahay RJ, Hodgson DJ, Boots M, Weber N et al. Using Social Network
493 Measures in Wildlife Disease Ecology, Epidemiology, and Management. *BioScience* 2017;
494 67(3):245–57. Available from: URL:
495 <https://academic.oup.com/bioscience/article/67/3/245/2962460>.

496 33. Silk MJ, Hodgson DJ, Rozins C, Croft DP, Delahay RJ, Boots M et al. Integrating social
497 behaviour, demography and disease dynamics in network models: applications to disease
498 management in declining wildlife populations. *Philos Trans R Soc Lond B Biol Sci* 2019;
499 374(1781):20180211.

500 34. Farine DR, Whitehead H. Constructing, conducting and interpreting animal social network
501 analysis. *Journal of Animal Ecology* 2015; 84(5):1144–63. Available from: URL:
502 <https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2656.12418>.

503 35. Hébert-Dufresne L, Young J-G, Bedson J, Skrip LA, Pedi D, Jalloh MF et al. The network
504 epidemiology of an Ebola epidemic; 2021 Nov 16. Available from: URL:
505 <https://arxiv.org/pdf/2111.08686>.

506 36. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998;
507 393(6684):440–2.

508 37. Rosillo N, Del-Águila-Mejía J, Rojas-Benedicto A, Guerrero-Vadillo M, Peñuelas M,
509 Mazagatos C et al. Real time surveillance of COVID-19 space and time clusters during the
510 summer 2020 in Spain. *BMC Public Health* 2021; 21(1):961.

511 38. Rozins C, Silk MJ, Croft DP, Delahay RJ, Hodgson DJ, McDonald RA et al. Social structure
512 contains epidemics and regulates individual roles in disease transmission in a group-living
513 mammal. *Ecology and Evolution* 2018; 8(23):12044–55. Available from: URL:
514 <https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.4664>.

515 39. Sah P, Leu ST, Cross PC, Hudson PJ, Bansal S. Unraveling the disease consequences and
516 mechanisms of modular structure in animal social networks. *PNAS* 2017; 114(16):4165–70.
517

Figures

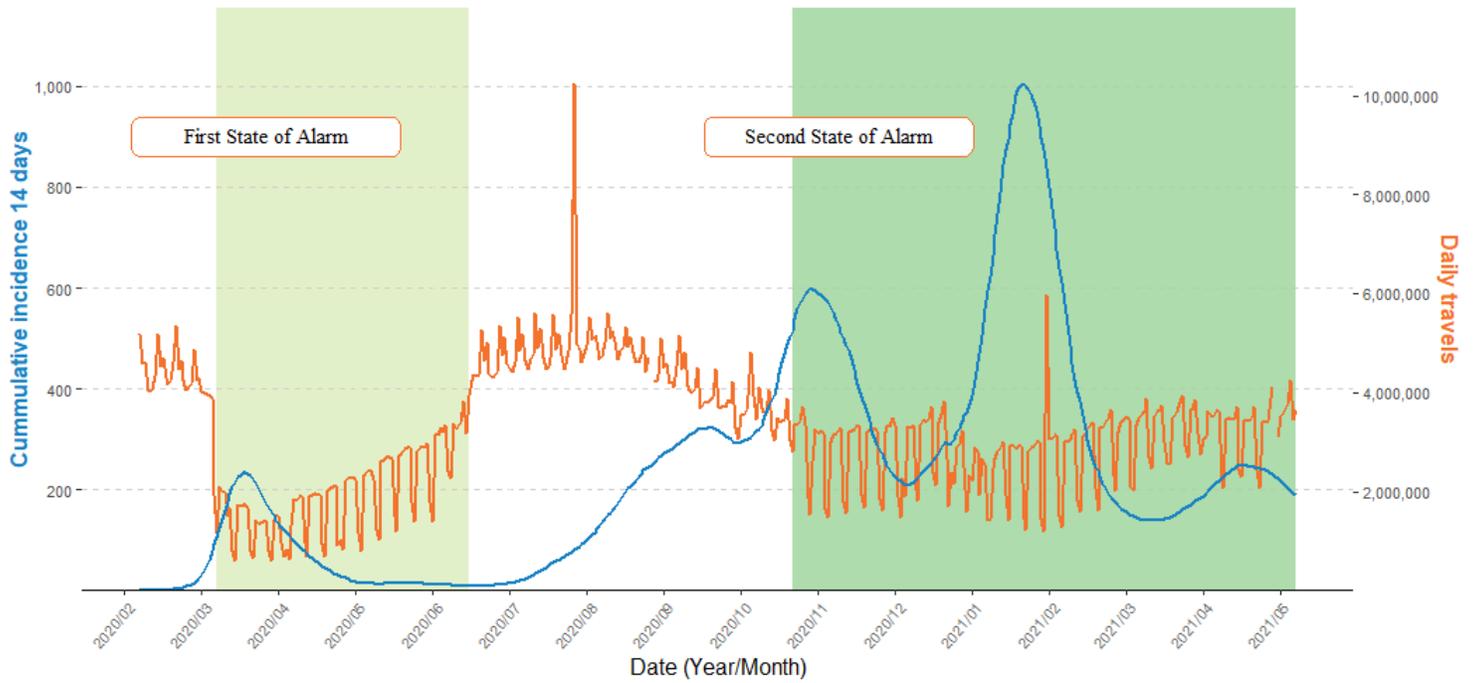
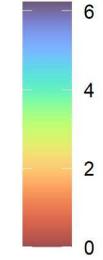


Figure 1

Daily total travels between provinces and 14-day cumulative incidence in Spain, 14th Feb 2020 – 9th May 2021

A

ED to Madrid



B

ED from Huesca

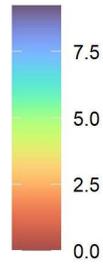


Figure 2

Shortest distances maps from (outgoing) and to (incoming) two provinces, Huesca and Madrid

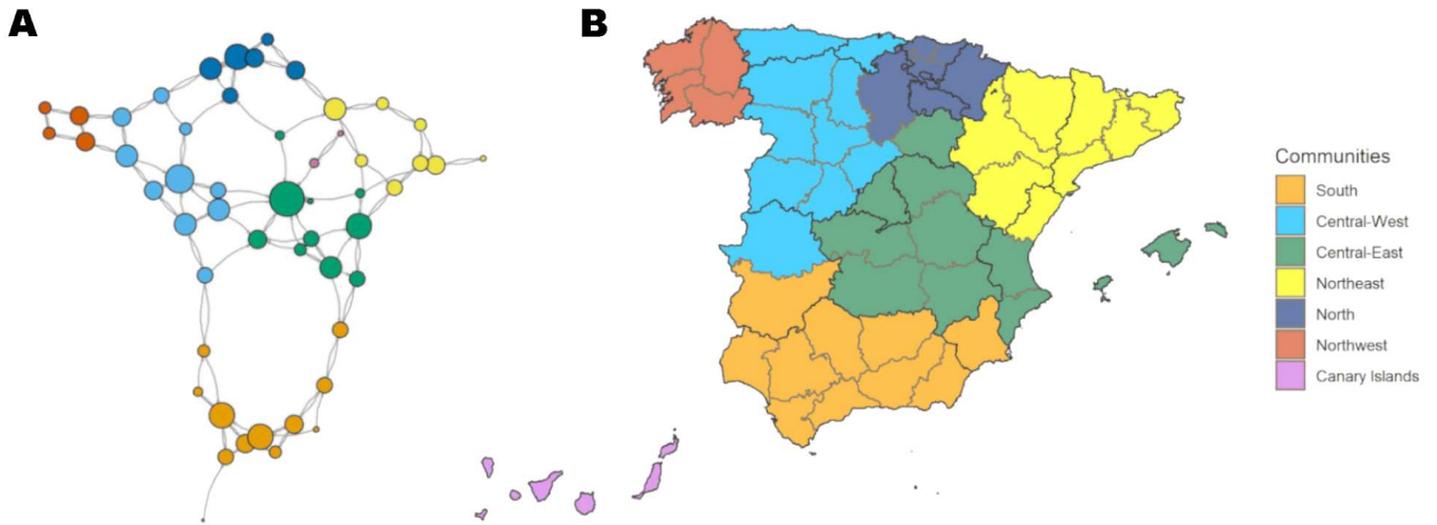


Figure 3

Effective distance network (EDN) representation of network measures and map of communities

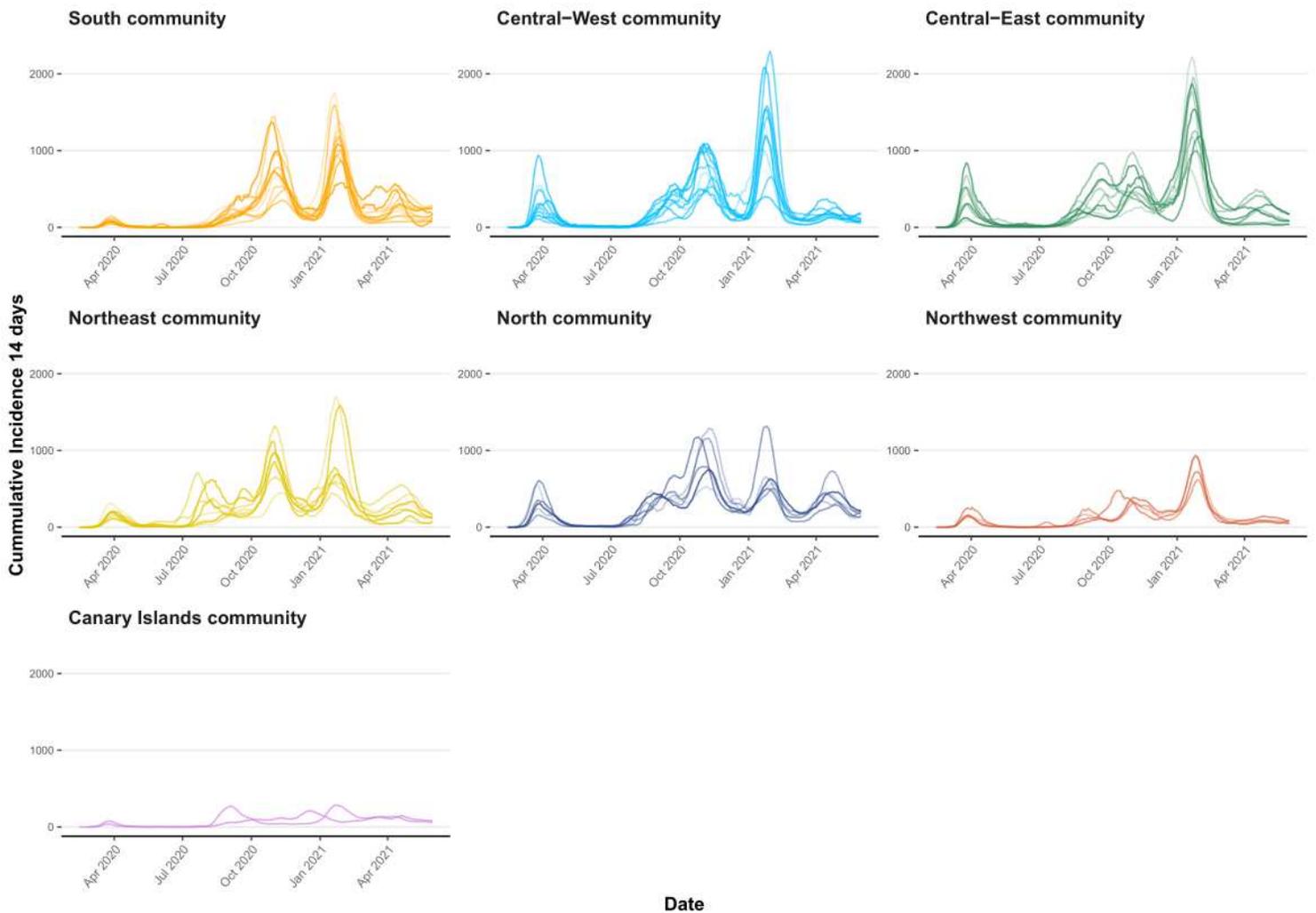


Figure 4

CI14 per 100.000 inhabitants by province, grouped by community during the study period

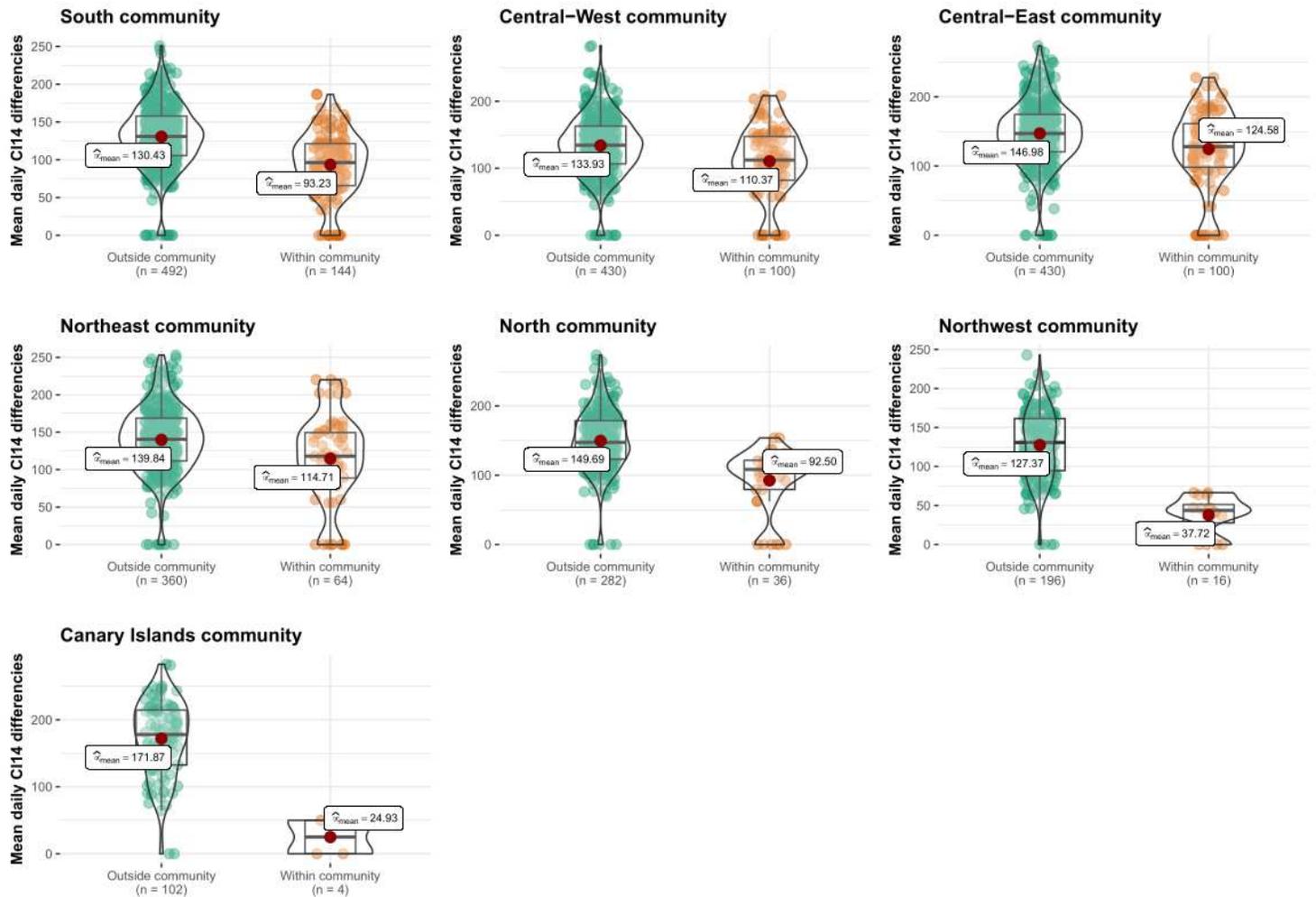


Figure 5

Mean daily differences between CI14 curves within communities' vs outside community for every detected community

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Map1.pdf](#)
- [Supplementaryfilefinal.docx](#)