



**HAL**  
open science

## **Cardiac dyspnea risk zones in the South of France identified by geo-pollution trends study**

Fanny Simões, Charles Bouveyron, Damien Piga, Damien Borel, Stéphane Descombes, Véronique Paquis-Flucklinger, Jaques Levraut, Pierre Gibelin, Silvia Bottini

### **► To cite this version:**

Fanny Simões, Charles Bouveyron, Damien Piga, Damien Borel, Stéphane Descombes, et al.. Cardiac dyspnea risk zones in the South of France identified by geo-pollution trends study. *Scientific Reports*, 2022, 12, pp.1900. <10.1038/s41598-022-05827-2>. <hal-03519068>

**HAL Id: hal-03519068**

**<https://hal.science/hal-03519068v1>**

Submitted on 10 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# 1 Cardiac dyspnea risk zones in the South of France

## 2 identified by geo-pollution trends study

3 Fanny Simões<sup>a</sup>, Charles Bouveyron<sup>b</sup>, Damien Piga<sup>c</sup>, Damien Borel<sup>d</sup>, Stéphane  
4 Descombes<sup>e</sup>, Véronique Paquis-Flucklinger<sup>f</sup>, Jaques Levraut<sup>\*g</sup>, Pierre Gibelin<sup>\*h</sup>,  
5 Silvia Bottini<sup>##a</sup>

6  
7 <sup>a</sup>Université Côte d'Azur, Center of Modeling, Simulation and Interactions, Nice, France;  
8 fanny.simoes@univ-cotedazur.fr

9 <sup>b</sup>Université Côte d'Azur, Inria, CNRS, Laboratoire JA Dieudonné, Maasai research team,  
10 Nice, France; charles.bouveyron@univ-cotedazur.fr

11 <sup>c</sup>ATMOSud, Nice, France ; damien.piga@atmosud.org

12 <sup>d</sup>Innovation e-Santé Sud, Hyères, France ; dborel@ies-sud.fr

13 <sup>e</sup>Université Côte d'Azur, Inria, CNRS, Laboratoire JA Dieudonné, Nice, France ;  
14 stephane.descombes@univ-cotedazur.fr

15 <sup>f</sup>Université Côte d'Azur, Inserm U1081, CNRS UMR7284, IRCAN, CHU de Nice, Nice,  
16 France ; veronique.paquis@univ-cotedazur.fr

17 <sup>g</sup>Département Hospitalo-Universitaire de Médecine d'Urgence, Nice, France ; levraut.j@chu-  
18 nice.fr

19 <sup>h</sup>Université Côte d'Azur, Faculté de médecine, Nice, France ; gibelin.p@chu-nice.fr

20 \* contributed as co-last authors

### 21 # Correspondence:

22 Silvia Bottini, PhD

23 Email : [silvia.bottini@univ-cotedazur.fr](mailto:silvia.bottini@univ-cotedazur.fr)

24 Tel: +33643670646

25 Address: 151 route de saint antoine de ginestiere, 06200 Nice (France).

### 27 Authors' emails :

28 Fanny Simoes; fanny.simoes@univ-cotedazur.fr

29 Charles Bouveyron; charles.bouveyron@univ-cotedazur.fr

30 Damien Piga ; damien.piga@atmosud.org

31 Damien Borel ; dborel@ies-sud.fr

32 Stephane Descombes ; stephane.descombes@univ-cotedazur.fr

33 Veronique Paquis-Flucklinger ; veronique.paquis@univ-cotedazur.fr

34 Jaques Levraut ; levraut.j@chu-nice.fr

35 Pierre Gibelin ; gibelin.p@chu-nice.fr

36 Silvia Bottini ; silvia.bottini@univ-cotedazur.fr

## 40 Abstract

41 **Background.** The incidence of cardiac dyspnea (CD) and the distribution of pollution in the  
42 south of France suggests that environmental pollution may have a role in disease triggering.  
43 CD is a hallmark symptom of heart failure leading to reduced ability to function and engage  
44 in activities of daily living.

45 **Methods.** To show the impact of short-term pollution exposure on the increment of CD  
46 emergency room visits, we collected pollutants and climate measurements on a daily basis  
47 and 43,400 events of CD in the Région Sud from 2013 to 2018. We used a distributed lag  
48 non-linear model (DLNM) to assess the association between air pollution and CD events.  
49 We divided the region in 357 zones to reconcile environmental and emergency room visits  
50 data. We applied the DLNM on the entire region, on zones grouped by pollution trends and  
51 on singular zones.

52 **Results.** Each pollutant has a significant effect on triggering CD. Depending on the pollutant,  
53 we identified four shapes of exposure curves to describe the impact of pollution on CD  
54 events: early and late effect for NO<sub>2</sub>; U-shape and rainbow-shape (or inverted U) for O<sub>3</sub>; all  
55 the four shapes for PM10. In the biggest cities, O<sub>3</sub> has the most significant association along  
56 with the PM10. In the west side, a delayed effect triggered by PM10 was found. Zones along  
57 the main highway are mostly affected by NO<sub>2</sub> pollution with an increase of the association for  
58 a period up to 9 days after the pollution peak.

59 **Conclusions.** Our results can be used by local authorities to set up specific prevention  
60 policies, public alerts that adapt to the different zones and support public health prediction-  
61 making. We developed a user-friendly web application called HEART (Health, Environment  
62 in PACA Region Tool) to collect our results. HEART will allow citizens, researchers and local  
63 authorities to monitor the impact of pollution trends on local public health.

64 **Key words**

65 Cardiac dyspnea; pollution; health – environment; distributed lag non-linear model; public  
66 health; delayed effects

67

## 68 **Background**

69 Cardiac dyspnea (CD) is a hallmark symptom of heart failure leading to reduced  
70 ability to function and engage in activities of daily living (1). This pathology  
71 particularly affects people aged more than 65 years old with a slight higher  
72 prevalence rate in women (2). Effective management of CD awaits a better  
73 understanding of its triggering features. Although cardiovascular factors are believed  
74 to play an important role, the experience of CD is multifactorial and may originate  
75 from different sources including environmental factors (3).

76 A growing body of evidence states that air pollution is a significant threat to health  
77 worldwide. The duration of the period of exposure to air pollution yields multiple  
78 consequences on health. A short-term exposure increases hospital admissions and  
79 mortality rate, causing mainly respiratory and cardiovascular diseases, including CD  
80 (4-5-6), whereas a long-term exposure reduces life expectancy (7-8-9-10-11).

81 Bourdrel et al. (12) showed that long-term exposure leads to an average increase of  
82 11% in cardiovascular mortality for an annual increase of  $10 \mu\text{g}/\text{m}^3$  in particulate  
83 matter less than 2.5 diameter (PM<sub>2.5</sub>). Increased hospitalizations and risk of  
84 mortality upon chronic exposure to PM<sub>2.5</sub> was also found in two epidemiological  
85 studies (13,14). Similarly, cardiovascular mortality increases for long-term and short-  
86 term exposures at dioxide nitrogen (NO<sub>2</sub>) (15). Ozone (O<sub>3</sub>) also has effects on health  
87 (16). Raza et al. (17) highlighted that short-term exposure to O<sub>3</sub> is associated with a  
88 high risk of cardiac arrest.

89 Although the relationship between short-term exposure to air pollution and several  
90 cardiovascular pathologies, such as acute myocardial infarction and congestive heart  
91 failure (18), is widely established; the link regarding CD is not yet fully demonstrated.  
92 The aim of this study is to shed a light on CD causation mechanisms and to develop  
93 prevention policies for health care by combining pollutants, environmental factors  
94 and CD hospitalization events.

95

## 96 **Methods**

### 97 **Study design**

98 The Région Sud (previously known as Provence-Alpes-Côte d'Azur, PACA region) is  
99 one of the eighteen administrative regions of France with a population of around 5  
100 millions of people covering a territory of 31,400 km<sup>2</sup>. It encloses six departments:  
101 Alpes-de-Haute-Provence (04), Alpes-Maritimes (06), Bouches-du-Rhône  
102 (13), Hautes-Alpes (05), Var (83) and Vaucluse (84) (figure 1A). The peculiarity of  
103 this region is that it presents a wide variety of landscapes: the Alps mountains, plains  
104 and coastal zones where the majority of the population lives. The region is subjected  
105 to different pollutants with extreme local specificity due to different distribution of  
106 industrial and agricultural activities, urbanization and traffic all over the territory.  
107 Furthermore, a demographic imbalance is also present, with 1 out of 4 citizens older  
108 than 60 years old. Altogether these characteristics make the Région Sud a suitable  
109 model to study the relationship between CD and pollution.

### 110 **Emergency room visits and environmental data**

111 Our study is based on two different databases. The emergency room visits data,  
112 provided by the agency ORUPACA (<https://ies-sud.fr/>) that consists of 43,400  
113 recorded events for individuals aged between 20 and 115 years old, who attended  
114 one of the 47 emergency centers of the Région Sud with symptoms related to CD  
115 between 2013 and 2018. Inclusion criteria include only patients with International  
116 Classification of Disease, version 10, (ICD-10,  
117 <https://icd.who.int/browse10/2019/en#/>) emergency codes: “I50” for heart failure and  
118 “I48” for atrial fibrillation and flutter (table 1). We decided to include atrial fibrillation  
119 and flutter and not acute coronary syndrome because an atrial dysrhythmia is a  
120 frequent cause of heart failure decompensation. On the other hand, acute coronary  
121 syndrome presentation is rarely an isolated dyspnea and the main clinical  
122 presentation is chest pain, even if ischemia might be the cause of cardiac  
123 decompensation. Pre-anonymized patient data are available, including age, gender  
124 and residential postal code (table 2).

125 Environmental data, provided by the agency AtmoSud (<https://www.atmosud.org/>),  
126 are composed by daily pollution and meteorological factors measurements over the  
127 period from 2013 to 2018. AtmoSud relies on a set of eighty static and mobile  
128 sensors which measure daily pollutants and meteorological variables. Regarding the  
129 pollutants, hourly measurements are taken during the day, but only the maximum  
130 value for  $NO_2$ , and  $O_3$ , and the average value for PM10 are registered. Then, using  
131 the sophisticated model Chimere (19) to interpolate sensors measurements, a  
132 detailed distribution map with a resolution of 4 km for each of these variables is  
133 released. Specifically, the region is divided in squares of 4 km, each square  
134 represents a measure, constituting a mesh of 1995 and 1967 squares for pollution  
135 and meteorological variables respectively. The environmental data are composed by

136 3 pollutants: the daily maximum observed of NO<sub>2</sub> and O<sub>3</sub> and the daily average value  
137 of PM<sub>10</sub>, and 4 meteorological factors, namely: the daily maximal observed  
138 temperature (T) and pressure (P) and the difference between the maximal and  
139 minimal value ( $\Delta T$  and  $\Delta P$ ).

## 140 Pre-processing and merging

### 141 Missing values and outliers of environmental data

142 The environmental database contained missing data, specifically there were 43 days  
143 missing (2%) for the pollution data and 22 (1%) for the meteorological data over  
144 2,191 days between 2013 and 2018. Outliers were also detected and were  
145 processed as missing data. An outlier is an abnormal value, namely a value that can  
146 impossibly be feasible for the associated variable. Two methods were used:  
147 interpolation and random sampling. The interpolation method was used for variables  
148 with a trend or seasonality (O<sub>3</sub>, PM<sub>10</sub>, P, T) and the random sampling method for  
149 variables with a random distribution (NO<sub>2</sub>), see Supplementary Information for  
150 details.

### 151 Merging emergency room visits and environmental data

152 The three databases had different resolutions: non-overlapping meshes of 4 km  
153 squares for pollution (figure 1B, insert I) and meteorological variables (figure 1B,  
154 insert II), and postal codes for emergency room visits data (figure 1B, insert III). In  
155 order to associate pollutant and meteorological measurements to each emergency  
156 room visit, we first needed to change the data resolution of the three databases.  
157 Thus, we created 357 zones (figure 1B, insert IV) of non-overlapping postal codes  
158 with associated a vector of 10 variables, namely: patient id, age, gender, maximal

159 value for  $\text{NO}_2$ ,  $\text{O}_3$ , T, P, daily temperature amplitude ( $\Delta T$ ) and daily pressure  
160 amplitude ( $\Delta P$ ) and the average value for PM10. Finally, we filtered out 8 zones for  
161 which no CD events were registered during the period under study, for a total of 349  
162 zones (figure 1B, insert V). Once the bases were reconciled, each emergency room  
163 visit has been associated to the environmental data of its residential zone. The  
164 database reconciliation procedure is fully explained in Supplementary Information.

## 165 Statistical analysis

166 We employed a distributed lag non-linear model (DLNM) (20). The DLNM allows to  
167 take into account the non-linear and delayed effects of environmental variables on  
168 CD events. It is usually employed for time series data. It is based on the definition of  
169 a « cross-basis » function, it describes simultaneously the shape of the relationship  
170 along both the space of the predictor and the lag dimension of its occurrence. We  
171 defined the cross-basis functions of environmental variables and their lags as natural  
172 cubic splines in order to take into account the trend and seasonality of the data. To  
173 choose the values of the parameters, we used as reference the zone corresponding  
174 to the city of Marseille because it has the highest number of CD events during the  
175 studied period. We tested several values of degree of freedom, between 2 and 5, for  
176 environmental variables and their lags, and 2 to 8 for time function. Among the  
177 meteorological variables, to avoid correlated variables, we included in our model only  
178 the maximum of amplitude of pressure and temperature and the maximum of  
179 temperature and pressure or the average of pressure and temperature. The  
180 selection was done according to the minimization of Q-AIC. The degree of freedom  
181 selected are: time function (2),  $\text{NO}_2$  (2), PM10 (3),  $\text{O}_3$  (3), T (2), P (4),  $\Delta T$  (2),  $\Delta P$  (2).  
182 The logarithm of the number of people for each zone (called here “population”) is

183 also included in the model. The number of habitants by INSEE code is found on  
184 website “*notreterritoire.maregionsud.fr*” and we calculated the population for our  
185 defined 357 zones. Since data were available only for the years up to 2016, we  
186 assumed the population to be constant for 2017 and 2018.

187 The association between environmental variables and CD events is studied on a  
188 period of 14 days (lags). The model applied is a generalized linear model with quasi-  
189 Poisson distribution. The explained variable is the number of CD by date and by  
190 zone. We represented the association between the outcome and predictors by the  
191 relative risk (RR) or overall effect for a specific increase above the threshold at each  
192 lag. Specifically, for each pollutant, we selected four levels of exposure from the  
193 World Health Organization (WHO) (<https://www.who.int/>) air quality guidelines,  
194 namely: “good”, “moderate”, “lightly”, “heavily” (intervals are indicated in table 3).  
195 Then we applied the statistical model based on DLNM to assess the RR for each of  
196 the four defined level when data are available for the studied period, compared to the  
197 basal reference value for each pollutant, for a period of 14 days before the event. For  
198 meteorological variables we divided the distribution in four percentiles (1<sup>st</sup>, 10<sup>th</sup>, 90<sup>th</sup>,  
199 99<sup>th</sup>) and the RR was assessed compared to median value (intervals are indicated in  
200 table 4). Since both high and low temperature can have an effect on the RR, to  
201 measure this association we chose the median because by definition is the value  
202 separating the higher half from the lower half of data distribution.

203 Singular pollutant model includes one pollutant and meteorological variables  
204 whereas multivariate pollutants model includes all pollutants and meteorological  
205 variables. In total, 8 models are tested (6 singular and 2 multivariate). The degree of  
206 freedom of pollutant variables ( $NO_2$ ,  $O_3$ , PM10) are fixed according to singular  
207 pollutant model whereas the degree of freedom of meteorological variables are fixed

208 according to multivariate pollutants model. Since singular and multivariate pollutants  
209 models gave similar results, only the last one is retained for analysis.

210

## 211 Multivariate meta-analysis

212 We employed a multivariate meta-analysis in order to synthesize results in multi-  
213 zones analysis with DLNM. It defines an average exposure-response association  
214 across the zones selected. The method of estimation for random effect multivariate  
215 meta-analysis used is restricted maximum likelihood (reml). The R package used is  
216 *mixmeta*.

217 This model was applied at three level on our data. Before running the model, we  
218 removed 97 zones having less than 18 CD events (minimal empirical value of  
219 employability of the model).

## 220 Global model on the entire region.

221 We performed a multivariate meta-analysis with the DLNM (21) on the previously  
222 defined zones in order to identify global trends on the risks of CD by pollutants at the  
223 region level (figure 1B, insert VI).

## 224 Global model on the zones grouped by pollution trends.

225 To take into account local fine structures of pollution distribution, we employed the  
226 co-clustering algorithm "*multiFunLBM*" (22) on the environmental dataset in order to  
227 identify groups of zones with homogeneous trends of pollutions (figure 1B, insert  
228 VII). To define groups of zones with similar trends of pollution, we used a specific co-  
229 clustering method developed for time series with R package *funLBM*. The algorithm  
230 takes into account 2 dimensions: space and time. The analysis is done by week of 7  
231 days (313 weeks for the 6 years). The aim is to identify zones and weeks having  
232 similar trends of environmental variables distributions. Parameters used are: the

233 basis function is Fourier, commonly used function for time series of periodic data,  
234 the type of initialization is funFEM. According to the Integrated Completed Likelihood  
235 (ICL) criterion, 6 clusters are obtained. The cluster 1 corresponds to countryside (69  
236 zones), cluster 2 to coast-side/medium size cities (23 zones), cluster 3 to mountains  
237 (31 zones), cluster 4 to coast-side/ big cities (17 zones), cluster 5 to highway (58  
238 zones) and cluster 6 to west side (53 zones). Finally, we applied the same model at  
239 the pollution clusters level.

240 **Local model on the singular zones.**

241 In order to identify specific local trends, we applied the DLNM singularly on each  
242 zone (figure 1B, insert VIII). Finally, we selected 23 zones with the highest number  
243 of CD events, representatives of each pollution cluster to explore the relationship  
244 between pollution and CD at the local level.

245

246 **Web application HEART**

247 In order to collect and to easily access to the results of this study, we developed a  
248 web application called HEART: Health – Environment PACA Region Tool. HEART  
249 can be easily downloaded at <https://github.com/UCA-MSI/HEART>.

250 **Results**

251 **Overall characteristics of emergency room visits data and**  
252 **environmental database**

253 The emergency room visits data is composed by 43,400 CD events collected by the  
254 47 emergency centers of Région Sud in the period from 1<sup>st</sup> January 2013 to 31<sup>st</sup>  
255 December 2018 reporting an event of CD with symptoms related to either heart  
256 failure (I50 codes, 98% of the events) or atrial fibrillation (I48 codes, 2% of the  
257 events) (table 1). A summary of emergency room visits data characteristic is  
258 reported in table 2.

259 We normalized the number of events by the total population by year and by zone  
260 and we reported the incidence of CD events in Région Sud in figure 2A. Indeed, only  
261 a small percentage of the population is concerned, slightly increasing over the period  
262 of the study. As expected, we observed a higher incidence of CD events in  
263 wintertime compared with summer (figure 2B) and no specific prevalence due to  
264 gender (figure 2C). The distribution of CD events by age follows the expected  
265 pattern: very few events among people younger than 70 years old with a narrow  
266 peak for people aged more than 90 years old (figure 2D). Finally, the distribution of  
267 CD events by department and by zone are reported in figure 2E and figure 2F,  
268 respectively. Overall similar percentages are observed, specifically the department of  
269 Alpes-de-Haute-Provence (04) located to the Alps reported the highest percentage  
270 of events (0.30%), along with the department of Var (83) on the coast (0.24%) while  
271 the lowest percentage is observed in the department of Bouches-du-Rhône (13)  
272 (0.14%).

273 The characteristics of environmental data are summarized in supplementary table 1  
274 and time series trends for each pollutant and meteorological variables in Région Sud  
275 in 2013 to 2018 are showed in Supplementary Figure 1. We observe that O<sub>3</sub> has a  
276 periodic behavior with peaks during summertime, PM10 has a less strong

277 seasonality and NO<sub>2</sub> does not show any periodicity. As expected, temperature has a  
278 seasonal pattern, while pressure shows a month periodic trend. No particular trend is  
279 reported for temperature and pressure amplitudes. The average values of each  
280 pollutant by zone, reported in Supplementary Figure 2, highlight the strong spatial  
281 distribution of each pollutant in the region. The correlations between air pollution and  
282 meteorological factors are presented in Supplementary Figure 3. The daily  
283 concentrations of O<sub>3</sub>, NO<sub>2</sub>, and PM10 are not correlated. Similarly, poor correlation  
284 between pollutant levels and meteorological factors is reported with exception of a  
285 good and positive correlation between O<sub>3</sub> and T (correlation coefficient  $r = 0.69$ ,  $P <$   
286  $0.05$ ), as expected. Finally, none of the environmental variable correlates with CD.  
287 The absence of correlation among the variables allow the use of the distributed lag  
288 non-linear model (DLNM) to identify the association between pollution and CD  
289 events.

## 290 Global model on the entire region finds significant 291 associations between meteorological factors and CD events.

292 We applied the DLNM model globally on the entire region and we measured the  
293 relative risk (RR) of having a CD event after exposure to pollution and meteorological  
294 factors (see methods for details on the model set-up). Results are reported by  
295 exposure curves to show global trends and by single day lag-response confidence  
296 intervals (CI) to show significance of RR.

297 Overall, no significant associations were found for NO<sub>2</sub> and PM10 (Supplementary  
298 figure 4A and 4C, respectively) on RR of CD when the model is applied globally on  
299 the entire region. Slightly significant associations are observed for O<sub>3</sub> when levels

300 raise above “heavily polluted” compared to reference level, increasing the RR up to  
301 1.06 (95% CI: 1-1.15), appearing at lag 0 and decreasing until no significant at lag 3  
302 (Supplementary figure 4B). These results suggest that the global model is not  
303 suitable for our data due to the different local distribution of pollution all over the  
304 region.

305 On the other side, regarding meteorological variables, the global model showed  
306 good performances, due to a more homogeneous distribution over the territory. As  
307 expected, low temperatures (1<sup>st</sup> percentile) increase the RR of CD progressively  
308 from 1.01 (95% CI: 1.00-1.05) at lag 3 up to 1.04 (95% CI: 1.02-1.06) at lag 14  
309 (figure 3A and Supplementary figure 4D) as well as high pressures (99<sup>th</sup> percentile)  
310 from lag 3 up to lag 12 (RR between 1.02 (95% CI: 1.00-1.04) and 1.04 (95% CI:  
311 1.01-1.06) (figure 3B and Supplementary figure 4E). High values of temperature  
312 amplitude (above the 99<sup>th</sup> percentile) increase RR of CD up to 1.04 (95% CI: 1.00-  
313 1.06) from lag 10 to lag 14 (figure 3C and Supplementary figure 4F). Finally, a slight  
314 association is observed only for high values of pressure amplitude (99<sup>th</sup> percentile) at  
315 lag 0 yielding a RR of 1.02 (95% CI: 1.00-1.04) up to lag 10 (figure 3D and  
316 Supplementary figure 4G). Overall, low temperature, high pressure and high  
317 amplitude values for these two parameters have a significant association on the RR  
318 of CD.

## 319 Global model by pollution clusters identified significant 320 associations between pollution and CD events

321 As already showed, the global model did not find any conclusive relationship  
322 between pollution exposure and CD events. We reasoned that the different territorial  
323 distribution of pollutants was not properly considered in the global analysis

324 (Supplementary figure 2). Thus, we took advantage of our previous work on the  
325 analysis of the environmental data in Région Sud (22), in which we grouped the 357  
326 zones in 6 clusters having homogeneous trends of pollution (figure 1B, insert VII and  
327 figure 4A), to exploit the impact of pollution on CD events. Curves of exposure and  
328 CI for significant RR values by lag for each cluster and pollutant are reported in  
329 figure 4 and supplementary table 2, respectively.

330 Exposure curves for NO<sub>2</sub> show a significant impact on RR only for cluster 6, that  
331 collects zones of the west side of the region, when the level overcomes the heavily  
332 polluted threshold, with linear trend increasing by lag (figure 4B).

333 For cluster 5, the association on RR becomes slightly significant with the increasing  
334 of pollution levels between lag 6 and lag 10 (figure 4C). An immediate association up  
335 to lag 2 is observed when O<sub>3</sub> level raises above lightly for cluster 1 (internal zones  
336 on the mountains) (figure 4D) and lightly+ for cluster 3 and up to lag 4 when O<sub>3</sub> level  
337 raises above good, moderate (figure 4E) .

338 PM10 has the highest impact on RR in cluster 4 with an immediate association at lag  
339 0 increasing significantly the RR up to 1.1 (95% CI: 1.00-1.3) regardless the level of  
340 pollution (figure 4F).

341 Overall, these results pointed out that RR of CD is affected by pollution levels, with  
342 trends and lags that depends by local territorial characteristics.

### 343 **Local model by singular zones identified specific exposure** 344 **curves trends**

345 Previous results, although significant, are mild. In order to exploit more how local  
346 pollution affects the RR of CD, we applied the DLNM statistical model singularly to

347 each of the 251 zones of Région Sud with at least 18 CD events. Then we selected  
348 23 zones with the highest number of CD events in order to guarantee a geographic  
349 representativity of the region peculiarities. Results for pollution variables on these  
350 zones are reported in figure 5, supplementary figure 5 and supplementary table 3.  
351 The study of several zones allowed the identification of common patterns of  
352 exposure curves characteristics of each pollutant. Therefore, four cumulative  
353 exposure curves shapes were identified: early and late effect for NO<sub>2</sub>; rainbow-shape  
354 (inverted U) and U-shape for O<sub>3</sub>; the four for PM<sub>10</sub> (figure 5, supplementary figure  
355 5).

356 NO<sub>2</sub> has a significant association with CD emergency room visits only in some zones  
357 belonging mainly to cluster 5, namely: “Grasse”, “Le Luc” (supplementary figure 5A),  
358 and mildly for “Salon de Provence” with two exposure patterns, namely “early effect”  
359 and “late effect” (figure 5A). This result is consistent with the characteristics of NO<sub>2</sub>  
360 pollution, mainly due to fuel combustion from motor vehicles and present in densely  
361 populated zones, and of the zones collected in these clusters passing by the main  
362 highway of the region. Notably, only for the zone “Nice” of cluster 4, NO<sub>2</sub> increases  
363 over the level heavily polluted yielding a RR of 1.08 (95% CI: 1.00-1.15) at lag 0  
364 (figure 5A). The presence of the airport along with densely populated area with  
365 elevated traffic might support this result.

366 Two main exposure patterns were observed for O<sub>3</sub> (figure 5B, supplementary figure  
367 5B): rainbow-shape (“Draguignan”, cluster 1; “Marseille”, “Aix en Provence” and  
368 “Toulon”, cluster 4; “Grasse”, cluster 5) or U-shape (“St. raphael”, cluster 2; “Salon  
369 de Provence”, cluster 5; “Avignon”, cluster 6). We observed that, often the rainbow-  
370 shape curves switch to U-shape trends when O<sub>3</sub> level increases over heavily  
371 polluted thresholds (“Draguignan”, cluster 1 and “Toulon” supplementary figure 5B,

372 cluster 4; “Marseille” , cluster 4, figure 5B). This switch could be due to an  
373 exacerbation of the association on the pathology going from delayed effect to  
374 immediate effect because the pollutant is not tolerated anymore by the individuals.  
375 This result suggests that the higher the O<sub>3</sub> pollution levels are, the earlier the  
376 association with CD events is significant. The zone “Hyères” in cluster 5 shows a late  
377 effect profile, with a significant association from lag 7 up to lag 14 when O<sub>3</sub> levels are  
378 higher than good, moderate and lightly polluted (figure 5B).

379 Several different exposure profiles were identified for PM<sub>10</sub> (figure 5C,  
380 supplementary figure 5C): early effect (“Marseille”, “Nice” and “Toulon”, cluster 4;  
381 “Salon de Provence”, cluster 5; “Avignon”, cluster 6), rainbow-shape (“Gap”, cluster  
382 1; “Cavaillon”, cluster 6), U-shape ( “La Farlède”, cluster 2; “La Ciotat”, cluster 5) and  
383 late effect (“Pertuis”, cluster 1; “Hyères”, cluster 5).

## 384 **Discussion and conclusions**

385 This is the first study of pollution associations on CD events done in Région Sud  
386 and, as far as we know, the first study to cover an entire region. We collected  
387 pollutants (NO<sub>2</sub>, O<sub>3</sub>, and PM<sub>10</sub>) and climate measurements (Temperature and  
388 Pressure) on a daily basis and 43,400 events of CD in the Région Sud from 2013 to  
389 2018. We present evidences that the three pollutants are associated with significant  
390 increment in CD emergency room visits. Four shapes of exposure curves were  
391 identified: early and late effect for NO<sub>2</sub>; rainbow-shape (inverted U) and U-shape for  
392 O<sub>3</sub>; all the four shapes for PM<sub>10</sub>. Importantly, our approach effectively proposes  
393 pollutants monitoring as tool for prediction-making for public health policies.

394 In this study, we have found significant associations between short-term exposure to  
395 NO<sub>2</sub>, O<sub>3</sub>, PM10 and CD events at the zone level. The strongest effect on CD  
396 incidence was observed for PM10, as for previous studies reported on other  
397 cardiovascular illnesses (23-24-25). This finding is supported by biological studies  
398 suggesting that particulate matter tends to deposit in pulmonary alveoli more than  
399 gaseous pollutants, triggering extra-pulmonary effects (13). Our study reported a  
400 positive association of CD events with O<sub>3</sub>, as already showed for acute coronary  
401 events (26). We observed a weak relationship between NO<sub>2</sub> and CD, as previously  
402 showed for acute myocardial infarction failure (18-23-24).

403 We found that not only extreme pollution levels could increase the risk of CD, but  
404 also moderate and good levels depending on the zone of the region. Importantly, the  
405 DLNM model used, depicted the non-linear and delayed influence of pollution  
406 exposure on CD events. Indeed, we observed four different shapes of exposure  
407 curves describing the relationship between pollutants and CD incidence, with  
408 important consequences for prevention and treatment. The early-effect curves  
409 suggest that the RR increases immediately after exposure and decreases after a  
410 certain period until no association on CD is observed. Contrarily, for late effect  
411 curves, the opposite scenario is to be expected. U-shaped relationship means that  
412 both immediate and late effects on the RR of CD events should be expected.  
413 Rainbow-shaped curves yield a slightly delayed effect on CD after the pollution peak,  
414 that persists for a certain period (depending on the amplitude of the significant lags).  
415 Accordingly, specific alerts can be set up by local authorities to prevent CD events.

416 One of the limitations of this study could be the resolution of the results: according to  
417 the General Data Protection Regulation, only patients' residential postal code was

418 available although data are anonymized. Thus, in order to associate pollution  
419 measurements to each emergency room visit, we had to change the resolution of  
420 pollution data from a mesh of 4 km squares to the zone level, yielding merged  
421 pollution values for zones covered by more than one square. This may have led an  
422 underestimation of exposure associations. However, since most of the zones  
423 matched with only one square, data aggregation is expected to play only a minor role  
424 in this underestimation. For the same reason, we cannot distinguish repeated  
425 emergency room visits from singular events. Multiple visits of the same individual  
426 would suggests an exacerbation of the pathology and might be affected by pollution  
427 differently. However the period covered by this study is relatively short to have  
428 several repeated visits for CD except for chronic CD that are excluded since they  
429 have a different ICD-10 code.

430 By contrast, some strengths should be discussed. First, the data on CD were  
431 collected from all the 47 emergency centers of the region allowing complete  
432 coverage and the sample size is quite big to achieve good statistical performances.  
433 Secondly, the availability of the exact date of CD events and daily measurements of  
434 pollutants allowed the temporal alignment between air pollution exposure and CD  
435 incidence. Thirdly, we set up a statistical model that can be applied at different  
436 scales, from the entire region up to single cities/zones.

437 In order to collect and to easily access to the results of this study, we developed a  
438 web application called HEART: Health – Environment PACA Region Tool. HEART  
439 can be used by local authorities in order to set up specific policies to lower down  
440 pollution or for public alerts when pollution raise above secure levels for the citizens  
441 that adapt to the different zones. Furthermore, it can help to identify periods of the

442 years that are particularly affected by pollution, in order to set up alerts and  
443 prevention behaviors for each specific zone.

444 In conclusion, the value of this study goes beyond the results obtained for the  
445 Région Sud, because the statistical model we set up, could be expanded to study  
446 other regions or zones and/or the impact of pollution on other pathologies.

## 447 **List of Abbreviations**

448 CD, cardiac dyspnea; RR, relative risk; DLNM, distributed lag non-linear model;  
449 NO<sub>2</sub>,dioxide nitrogen; O<sub>3</sub>,Ozone; PM, particulate matter.

450

## 451 **Declarations**

### 452 **Data availability**

- 453 • Project name: HEART
- 454 • Project home page: <https://github.com/UCA-MSI/HEART>
- 455 • Operating system(s): Platform independent
- 456 • Programming language: R
- 457 • Other requirements: R 3.6 or higher.
- 458 • License: e.g. GNU GPL 3.0.
- 459 • Any restrictions to use by non-academics: none.

460 The authors declare that the data supporting the findings of this study are available  
461 within the article and its supplementary information files.

### 462 **Competing interests**

463 The authors declare that they have no competing interests.

## 464 Funding

465 This work was supported by the French government, through the UCA<sup>JEDI</sup>  
466 Investments in the Future project and 3IA Cote d'Azur managed by the National  
467 Research Agency (ANR) under reference numbers ANR-15-IDEX-01 and ANR-19-  
468 P3IA-0002, respectively.

## 469 Authors' Contributions

470 Conceptualization: SB, CB, PG, JL. Data curation: DP, DB, FS. Formal analysis: FS.  
471 Investigation: FS. Methodology: SB. Project administration: SB, VP. Resources: DP,  
472 DB. Supervision: SB, CB, PG, JL, SD. Validation: SB, CB. Visualization: FS.  
473 Software: FS. Writing – original draft: SB. Writing – review & editing: SB, VP, PG.

474

## 475 References

- 476 1. Cardiology ES of. The ESC Textbook of Cardiovascular Medicine. OUP Oxford; 2009.  
477 1423 p.  
478
- 479 2. van Mourik Y, Rutten FH, Moons KGM, Bertens LCM, Hoes AW, Reitsma JB.  
480 Prevalence and underlying causes of dyspnoea in older people: a systematic review. *Age*  
481 *Ageing*. 2014;43(3):319-26.  
482
- 483 3. Peters SP. When the Chief Complaint Is (or Should Be) Dyspnea in Adults. *The Journal of*  
484 *Allergy and Clinical Immunology: In Practice*. 2013;1(2):129-36.  
485
- 486 4. Benbrahim-Tallaa L, Baan RA, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V,  
487 et al. Carcinogenicity of diesel-engine and gasoline-engine exhausts and some nitroarenes.  
488 *The Lancet Oncology*. 2012;13(7):663-4.  
489
- 490 5. Hamra Ghassan B., Guha Neela, Cohen Aaron, Laden Francine, Raaschou-Nielsen Ole,  
491 Samet Jonathan M., et al. Outdoor Particulate Matter Exposure and Lung Cancer: A  
492 Systematic Review and Meta-Analysis. *Environmental Health Perspectives*.  
493 2014;122(9):906-11.  
494

- 495 6. Lavigne E, Burnett R, Weichenthal S. Association of short-term exposure to fine  
496 particulate air pollution and mortality: effect modification by oxidant gases. *Scientific*  
497 *Reports*. 2018;8.  
498
- 499 7. Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air  
500 pollution sources to premature mortality on a global scale. *Nature*. 2015;525(7569):367-71.  
501
- 502 8. Pascal M, de Crouy Chanel P, Wagner V, Corso M, Tillier C, Bentayeb M, et al. The  
503 mortality impacts of fine particles in France. *Science of The Total Environment*.  
504 2016;571:416-25.  
505
- 506 9. Hadley MB, Vedanthan R, Fuster V. Air pollution and cardiovascular disease: a window of  
507 opportunity. *Nat Rev Cardiol*. 2018;15(4):193-4.  
508
- 509 10. Rajagopalan S, Al-Kindi SG, Brook RD. Air Pollution and Cardiovascular Disease:  
510 JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2018;72(17):2054-70.  
511
- 512 11. Lelieveld J, Klingmüller K, Pozzer A, Pöschl U, Fnais M, Daiber A, et al. Cardiovascular  
513 disease burden from ambient air pollution in Europe reassessed using novel hazard ratio  
514 functions. *European Heart Journal*. 2019;40(20):1590-6.  
515
- 516 12. Bourdrel T, Bind M-A, Béjot Y, Morel O, Argacha J-F. Cardiovascular effects of air  
517 pollution. *Archives of Cardiovascular Diseases*. 2017;110(11):634-42.  
518
- 519 13. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al.  
520 Particulate matter air pollution and cardiovascular disease: An update to the scientific  
521 statement from the American Heart Association. *Circulation*. 2010;121(21):2331-78.  
522
- 523 14. Chen H, Burnett RT, Kwong JC, Villeneuve PJ, Goldberg MS, Brook RD, et al. Spatial  
524 association between ambient fine particulate matter and incident hypertension. *Circulation*.  
525 2014;129(5):562-9.  
526
- 527 15. Atkinson RW, Butland BK, Anderson HR, Maynard RL. Long-term Concentrations of  
528 Nitrogen Dioxide and Mortality: A Meta-analysis of Cohort Studies. *Epidemiology*.  
529 2018;29(4):460-72.  
530
- 531 16. Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. Ambient  
532 PM<sub>2.5</sub>, O<sub>3</sub>, and NO<sub>2</sub> Exposures and Associations with Mortality over 16 Years of Follow-  
533 Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health*  
534 *Perspect*. 2015;123(11):1180-6.  
535
- 536 17. Raza A, Dahlquist M, Jonsson M, Hollenberg J, Svensson L, Lind T, et al. Ozone and  
537 cardiac arrest: The role of previous hospitalizations. *Environmental Pollution*. 2019;245:1-8.  
538
- 539 18. Bai L, Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, et al. Exposure to  
540 ambient air pollution and the incidence of congestive heart failure and acute myocardial  
541 infarction: A population-based study of 5.1 million Canadian adults living in Ontario.  
542 *Environment International*. 2019;132:105004.  
543

- 544 19. Menut L, Bessagnet B, Khvorostyanov D, Beekmann M, Blond N, Colette A, et al.  
545 CHIMERE 2013: a model for regional atmospheric composition modelling. *Geoscientific*  
546 *Model Development*. 2013;6(4):981-1028.  
547
- 548 20. Gasparri A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in*  
549 *Medicine*. 2010;29(21):2224-34.  
550
- 551 21. Gasparri A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and  
552 other multi-parameter associations. *Statistics in Medicine*. 2012;31(29):3821-39.  
553
- 554 22. Bouveyron C, Jacques J, Schmutz A, Simoes F, Bottini S. Co-Clustering of Multivariate  
555 Functional Data for the Analysis of Air Pollution in the South of France. 2020.  
556
- 557 23. Cesaroni G, Forastiere F, Stafoggia M, Andersen ZJ, Badaloni C, Beelen R, et al. Long  
558 term exposure to ambient air pollution and incidence of acute coronary events: prospective  
559 cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. *BMJ*.  
560 2014;348:f7412.  
561
- 562 24. Lipsett MJ, Ostro BD, Reynolds P, Goldberg D, Hertz A, Jerrett M, et al. Long-term  
563 exposure to air pollution and cardiorespiratory disease in the California teachers study cohort.  
564 *Am J Respir Crit Care Med*. 2011;184(7):828-35.  
565
- 566 25. Puett RC, Hart JE, Yanosky JD, Paciorek C, Schwartz J, Suh H, et al. Chronic fine and  
567 coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health  
568 Study. *Environ Health Perspect*. 2009;117(11):1697-701.  
569
- 570 26. Ruidavets Jean-Bernard, Cournot Maxime, Cassadou Sylvie, Giroux Michel, Meybeck  
571 Mariam, Ferrières Jean. Ozone Air Pollution Is Associated With Acute Myocardial  
572 Infarction. *Circulation*. 2005;111(5):563-9.  
573

574  
575  
576  
577  
578  
579  
580  
581  
582

## 583 **Figure Legends**

584

585 **Figure 1.** Study design and data analysis workflow. A) The Région Sud and its 6  
586 departments.

587 The three databases used had different resolutions: non-overlapping meshes of 4  
588 km squares for pollution (B I) and meteorological variables (B II), and postal codes

589 for emergency room visits data (B III). To reconcile the databases, we created 357

590 zones (B IV) of non-overlapping postal codes. 8 zones, for which no CD events were  
591 registered during the period under study, were filtered out, for a total of 349 zones (B  
592 V). Finally, the model was run globally on 251 zones out of 349 because 97 zones  
593 contain less than 18 CD events (minimal empirical value of employability of the  
594 model) over the entire period of 6 years concerned by this study (B VI). Then the  
595 model was run on zones grouped by pollution trends (B VII). Finally, the model was  
596 run on the 251 zones singularly and 23 zones with the highest number of CD events  
597 were selected (B VIII). Maps were created with the R package mapprools v1.1-1.

598

599 **Figure 2.** Main characteristics of emergency room visits data. The incidence of CD  
600 events in Région Sud (A), by season (B) and by gender (C). The distribution of CD  
601 events by age (D) by department (E) and by zone (F). Maps were created with the R  
602 package mapprools v1.1-1.

603

604 **Figure 3.** Global DLNM (only significant results are reported). Lag-specific effect on  
605 CD events for increases of each of the meteorological variables temperature (A),  
606 pressure (B), temperature amplitude (C) and pressure amplitude (D), over the  
607 thresholds indicated in the legend in the plot for all zones in Région Sud.

608

609 **Figure 4.** DLNM by pollution clusters (only significant results are reported). A) Zones  
610 of Région Sud coloured by cluster of pollution. Colour code: light blue, cluster 1  
611 (internal zones on the mountains); light green, cluster 3 (mountains); dark green,  
612 cluster 4 (big cities on the coast); pink, cluster 5 (highway); red, cluster 6 (west  
613 side); grey, zones not included. B) Lag-specific effect on CD events for increases of  
614 NO<sub>2</sub> levels over the two thresholds as indicated in the legend in the plot for zones in  
615 cluster 6 (west side). Lag-specific effect on CD events for increases of O<sub>3</sub> levels  
616 over the thresholds indicated in the legend in the plot for zones in cluster 5 (highway)  
617 (C), cluster 1 (internal zones on the mountains) (D) and cluster 3 (mountains) (E). F)  
618 Lag-specific effect on CD events for increases of PM<sub>10</sub> over the thresholds indicated  
619 in the legend in the plot for zones in cluster 4 (big cities on the coast). Maps were  
620 created with the R package mapprools v1.1-1.

621

622

623 **Figure 5.** DLNM by singular zones, only significant results for the selected 23 zones  
624 are reported. Lag-specific effect on CD events for increases of NO<sub>2</sub> (A), O<sub>3</sub> (B) and  
625 PM10 (C) over the thresholds indicated in the legend in the plot. Only the most  
626 representative plot is reported for each pollutant and shape of the response-effect  
627 curves. Others zones showing similar trends are reported in Supplementary Figure  
628 5. Maps were created with the R package mapproj v1.1-1.

629  
630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671

## Tables

**Table 1.** Inclusion codes for patient population.

Emergency code	Explication	Number of events
I50=I50.0, I50.1, I50.9, I500, I501, I509	Heart failure	42,769
I48 = I48.0, I48.1, I48.2, I48.3, I48.4, I48.9 associated with codes: J18.9, T17.9, R06.0, T75.1	Atrial fibrillation and flutter	631

675  
676  
677  
678  
679  
680  
681

682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702

**Table 2.** Main characteristics of the study population in Région Sud from 2013 to 2018.

Characteristics	Total	2013	2014	2015	2016	2017	2018
Gender n (%)							
Male	19 991 (46)	2 660 (45)	2862 (46)	3 089 (46)	3 373 (46)	3 926 (46)	4 081 (47)
Female	23 409 (54)	3 261 (55)	3 406 (54)	3 575 (54)	3 933 (54)	4 542 (54)	4 692 (53)
Age mean	82	82	82	82	82	82	82
Age group n (%)							
20-59	1 784 (4)	250 (4)	258 (4)	302 (4)	308 (4)	339 (4)	327 (4)
60-69	3 468	459	502	519	585	695	708

	(8)	(8)	(8)	(8)	(8)	(8)	(8)
70-79	7 984 (18)	1123 (19)	1 171 (19)	1 193 (18)	1 325 (18)	1 535 (18)	1 637 (19)
80-89	18 947 (44)	2661 (45)	2 837 (45)	2 933 (44)	3 269 (45)	3 610 (43)	3 637 (41)
>90	11 217 (26)	1428 (24)	1 500 (24)	1 717 (26)	1 819 (25)	2 289 (27)	2 464 (28)
Department n (%)							
04	2 300 (5)	336 (6)	338 (6)	351 (5)	380 (5)	397 (5)	498 (6)
05	1 196 (3)	227 (4)	193 (3)	170 (3)	200 (3)	197 (2)	209 (2)
06	9 608 (22)	1 214 (21)	1 339 (21)	1 486 (22)	1 674 (23)	1 824 (21)	2 071 (24)
13	12 501 (29)	1 628 (27)	1 772 (28)	1 770 (27)	2 108 (29)	2 589 (31)	2 634 (30)
83	11 968 (28)	1 786 (30)	1 743 (28)	1 985 (30)	1 961 (27)	2 294 (27)	2 199 (25)
84	5 827 (13)	730 (12)	883 (14)	902 (13)	983 (13)	1 167 (14)	1 162 (13)

703

704 **Table 3.** Pollutant threshold ranges associated to selected levels for this study from  
705 the WHO guidelines.

Levels of pollution	NO <sub>2</sub> (µg/m <sup>3</sup> )	O <sub>3</sub> (µg/m <sup>3</sup> )	PM10 (µg/m <sup>3</sup> )
Reference (very good)	50	40	10
Good	61-120	55-108	16-30
Moderate	121-140	109-126	31-35
Lightly	141-200	127-180	36-50

Heavily	>200	>180	>50
---------	------	------	-----

706

707

708

709

710

711