Projekt 2

Analiza zmienności zanieczyszczenia powietrza w Krakowie pyłami PM 2.5

Sprawozdanie "Metody komputerowe w badaniach klimatu i środowiska" Rok akademicki 2020/2021, AGH, Geoinformatyka, III rok

> Natalia Gadocha, nr 304165 Katarzyna Kowalczyk, nr 304174 Karolina Kuna, nr 304177

Wstęp	3
Cel projektu	3
Problematyka	3
Rodzaje zanieczyszczeń	3
Przebieg projektu	4
Dane	4
Przygotowanie danych	5
Podział wyników	8
Wyniki - godziny poranne i nocne	8
Zanieczyszczenia	9
Temperatura	10
Wilgotność	12
Błędy pomiarowe	14
Analiza dla biegaczy - godziny wieczorne	14
Zanieczyszczenia	14
Temperatura	19
Wilgotność	24
Wnioski	29
Godziny poranne i nocne	29
Analiza dla biegaczy	30
Wykorzystane procedury	31
Biblioteki	31
Procedury	33

1. Wstęp

1.1. Cel projektu

Celem projektu będzie zbadanie, jak przedstawia się zanieczyszczenie powietrza w Krakowie w różnych porach dnia.

Pierwsza część to analiza zanieczyszczeń, temperatury i wilgotności dla poranka i późnego wieczoru. Druga część zawiera analizę powyższych parametrów dla osób aktywnych - szczególnie tych, którzy preferują wieczorne spacery oraz bieganie.

1.2. Problematyka

Problem zanieczyszczenia powietrza jest jednym z poważniejszych zagadnień, dotyczących bezpośredniego wpływu przemysłu na człowieka i środowisko. Według raportu szwajcarskiej organizacji IQAir, ponad 90% społeczeństwa oddycha aktualnie niebezpiecznie skażonym powietrzem. Najnowszy raport powyższej platformy monitoringu donosi także, że aż 29% najbardziej zanieczyszczonych europejskich miast znajduje się w Polsce. Kraków uplasował się na 7. pozycji. Jesteśmy zatem jednym z państw, których problem dotyka najbardziej.

Problem zanieczyszczonego powietrza w dużej mierze dotyka biegaczy. Ich regularna, często codzienna aktywność powoduje dużą ekspozycję na groźne dla zdrowia pyły 2.5 PM i inne szkodliwe związki. Postanowiliśmy więc zbadać szczegółowo powietrze w godzinach wieczornych. Jest to pora, kiedy najwięcej sportowców odbywa treningi. Przedstawimy statystyki dotyczące zanieczyszczenia i temperatury. Porównamy ze sobą także wilgotność - ma ona duże znaczenie szczególnie dla biegaczy trenujących profesjonalnie. Im jest wyższa, tym proces parowania jest wyraźnie ograniczony. Ciepło jest kumulowane w organizmie, powodując systematyczny wzrost temperatury ciała, co może prowadzić do stanu przegrzania i odwodnienia.

1.3. Rodzaje zanieczyszczeń

Spośród zanieczyszczeń zawartych w powietrzu możemy wyróżnić przede wszystkim tlenki azotu, tlenki węgla i pyły zawieszone. Według danych Światowej Organizacji Zdrowia (WHO) najgroźniejszym z wymienionych składników są pyły 2.5 PM, które w znaczący sposób wpływają na zdrowie człowieka. Raport Europejskiej Agencji Środowiska (EEA) z 2020 roku wskazuje zanieczyszczenie powietrza jako główną przyczynę przedwczesnej

śmierci - ok. 400 tysięcy osób rocznie traci z tego powodu życie. Najczęstsze przypadki śmierci to choroby serca i zawały, a także nowotwory płuc, wywoływane zbyt wysokim stężeniem 2.5 PM.

2. Przebieg projektu

2.1. Dane

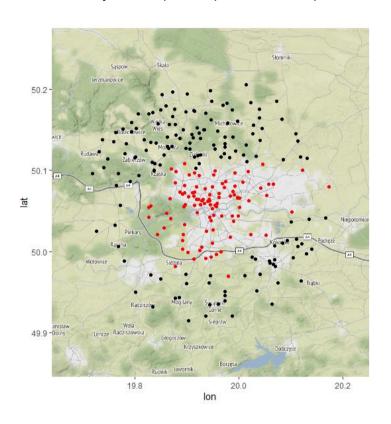
W projekcie wykorzystano dane, zgromadzone w serwisie internetowym www.airly.org. Są to dane dotyczące zanieczyszczenia powietrza pyłami 2.5PM, temperatury oraz wilgotności. Informacje zebrane zostały z kilkudziesięciu punktów pomiarowych na terenie całego Krakowa, w dniach 16.01.2021 (sobota), 17.01.2021 (niedziela) oraz 18.01.2021 (poniedziałek). Po wczytaniu danych do programu RSTudio, wykonano szereg czynności, prowadzących do utworzenia mapy. Do estymacji brakujących danych wykorzystano metodę krigingu.

Kolejne dane pozyskane zostały z serwisu Google Maps: https://www.google.com/maps. Dotyczą one natężenia ruchu w Krakowie o określonych porach dnia - podobnie jak dane opisane powyżej.

2.2. Przygotowanie danych

Korzystając z programu RStudio przygotowano mapki pomocnicze.

1. Rozmieszczenie czujników w postaci punktów na mapce:

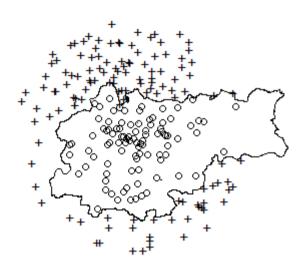


2. Dzielnice Krakowa, dla których pozyskane zostały dane:



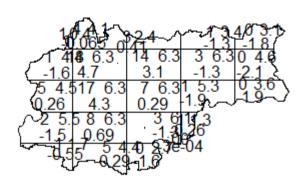
3. Rozkład czujników. Okręgi to punkty na terenie Krakowa, krzyżyki - poza granicami miasta:

data15_ppp



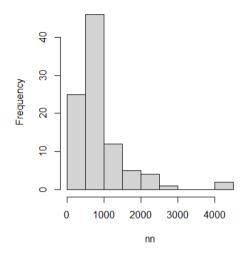
4. Statystyki rozkładu czujników:

qt

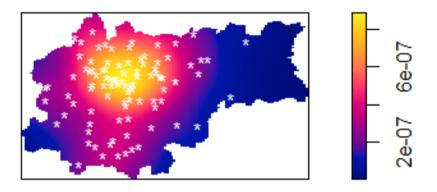


P-value wynosi 4.325e-08, co odrzuca hipotezę o losowości przestrzennej. Innym sposobem na sprawdzenie tezy jest wykonanie histogramu lub mapy intensywności, które przedstawione zostaną w kolejnych punktach.

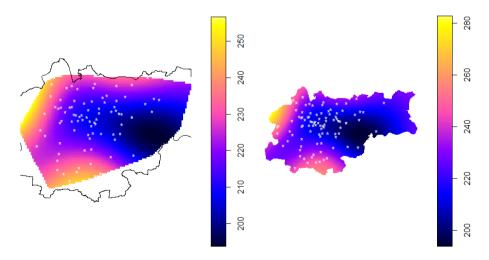
5. Histogram:



6. Mapa intensywności:



7. Pierwsze próby przedstawienia danych:



Mapa została przedstawiona na dwa sposoby.

Drugi sposób obejmował dużo obszarów bez czujników, gdzie błąd pomiarowy miał wysokie wartości. Mapa pierwsza pozwalała za to na większe zbliżenie w najbardziej interesujące nas obszary, gdzie czujników jest dużo, dlatego właśnie ona została wybrana na obowiązującą w projekcie.

2.3. Podział wyników

Wyniki podzielono na dwie kategorie, opisane poniżej. Każda z nich zawiera zebrane dane, obserwacje oraz wnioski.

Godziny poranne i nocne

Pierwszą częścią projektu była analiza danych zebranych w nocy i rano.

Dane pobierane były o północy oraz o godz. 6:00.

Są to pory dnia, kiedy zazwyczaj ruch na ulicach jest najmniejszy - zbadamy, czy rzeczywiście przekłada się to na mniejsze zanieczyszczenie pyłem 2.5PM

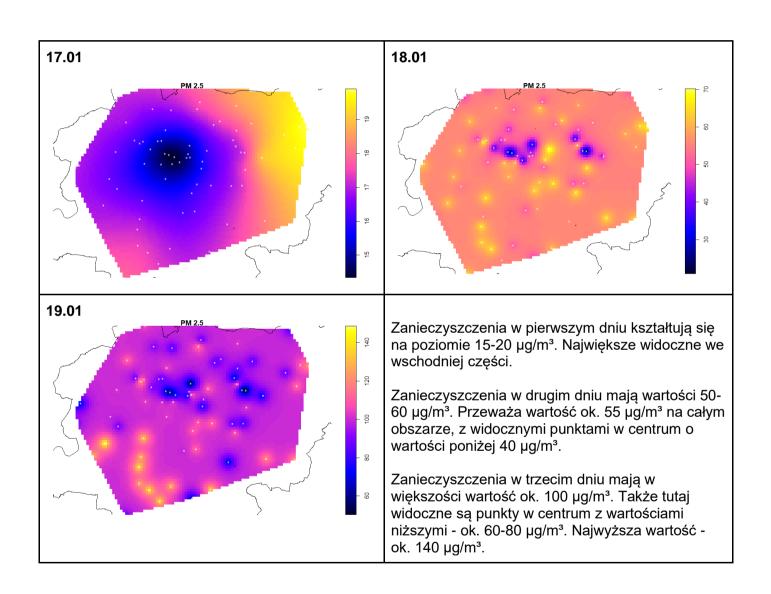
Dla aktywnych

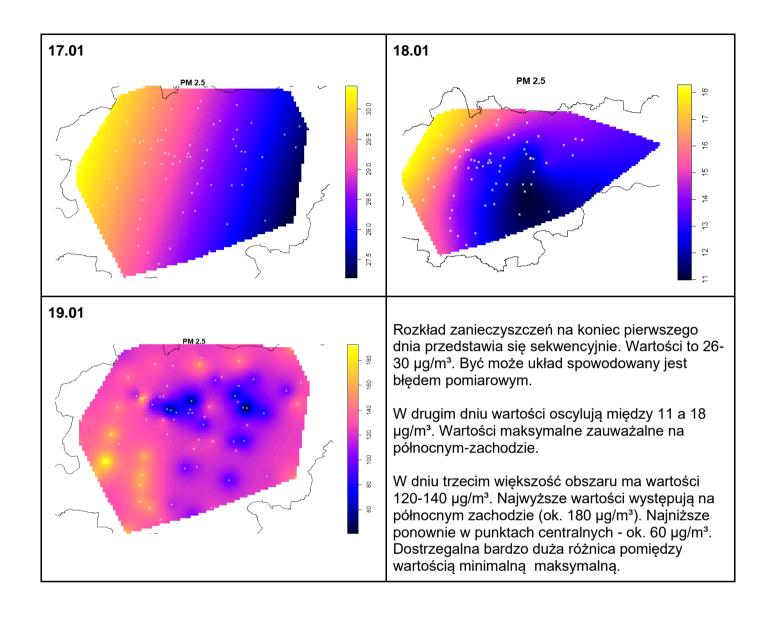
Drugą część projektu stanowiła analiza godzin wieczornych. Zestawienie dla aktywnych zostało przygotowane z myślą o biegaczach i osobach, które wieczorem spędzają czas na świeżym powietrzu. Do grupy odbiorców zaliczyć można także spacerowiczów z psami i każdego, kto wychodzi wieczorem na zewnątrz. Dane zostały pobrane kolejno o godzinach: 17:00, 18:00, 19:00, 20:00, 21:00. Zazwyczaj jest to pora czasu wolnego po pracy, kiedy dużo osób wychodzi w celach sportoworekreacyjnych.

3. Wyniki - godziny poranne i nocne

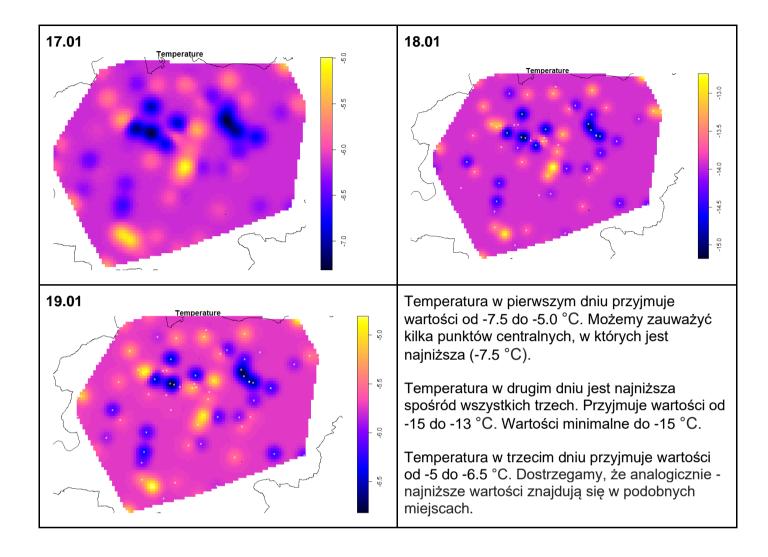
3.1. Zanieczyszczenia

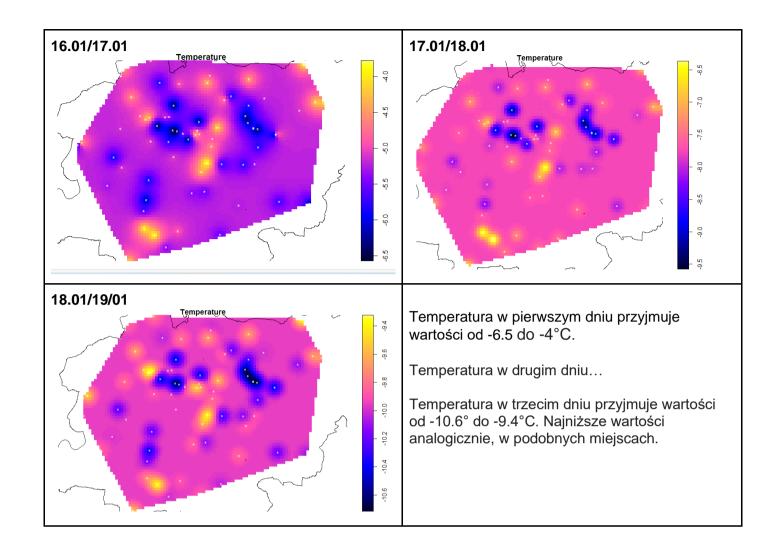
GODZ. 6:00



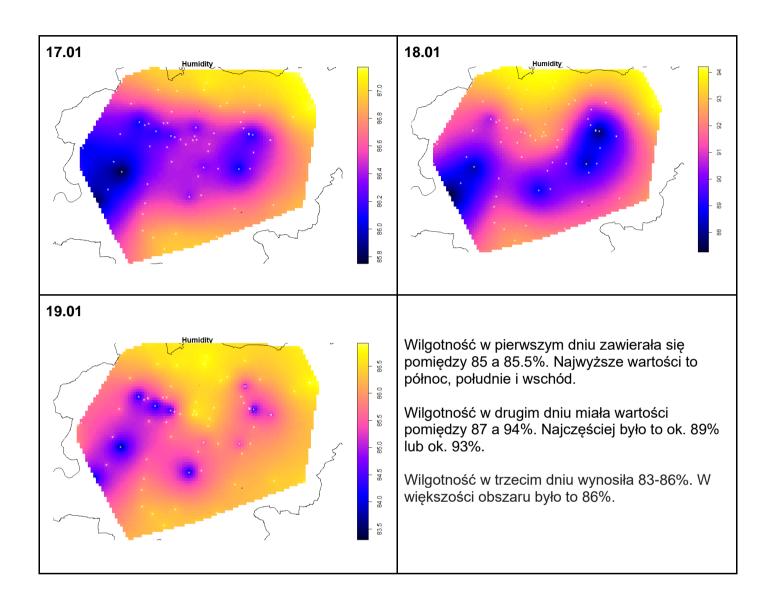


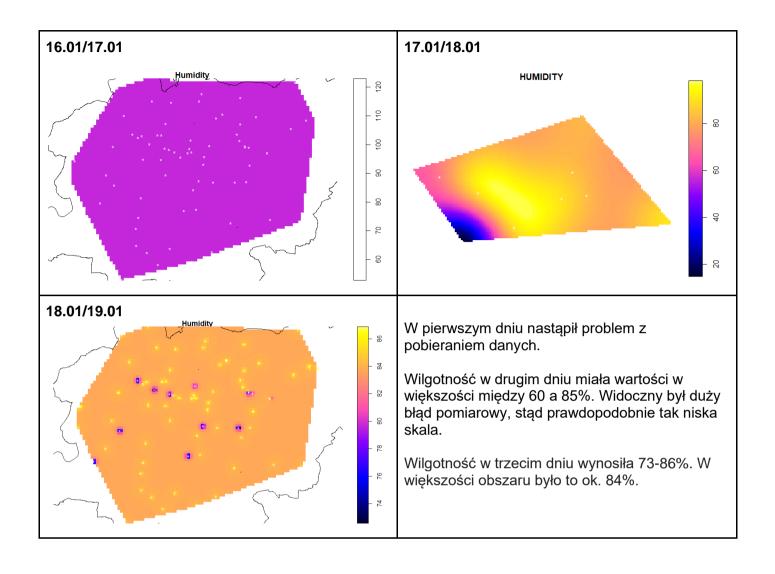
3.2. Temperatura





GODZ. 6:00

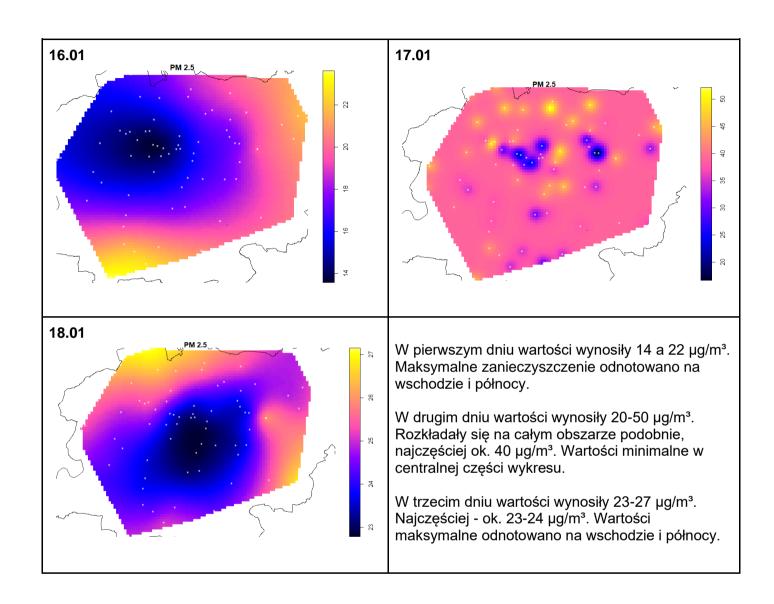




4. Analiza dla biegaczy - godziny wieczorne

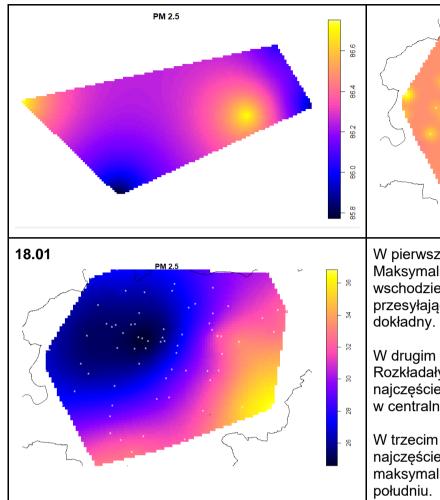
4.1. Zanieczyszczenia

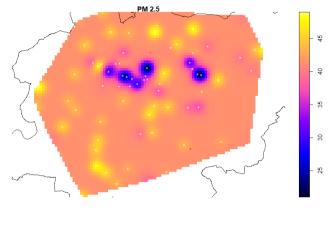
GODZ. 17:00



GODZ. 18:00

16.01	17.01
-------	-------

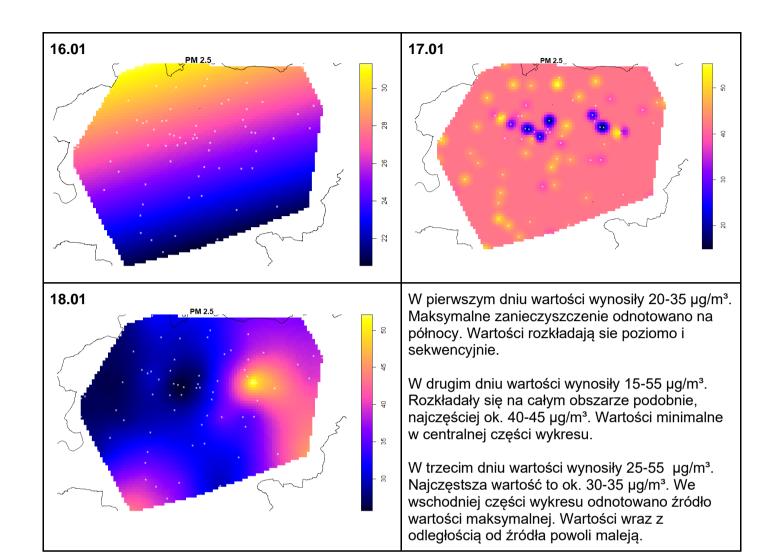


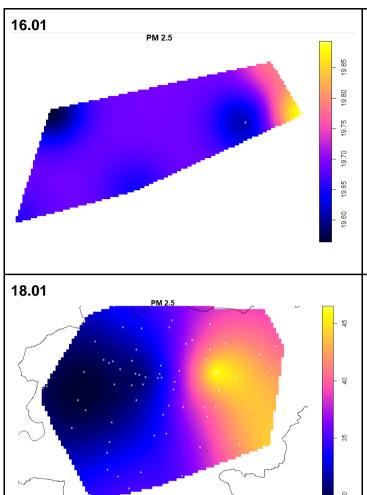


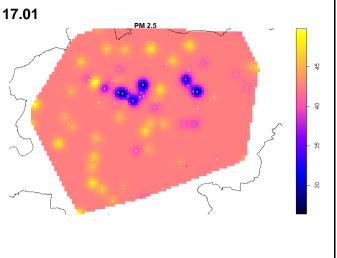
W pierwszym dniu wartości wynosiły 77-85 μg/m³. Maksymalne zanieczyszczenie odnotowano na wschodzie. Z powodu dużej liczby czujników przesyłających brak informacji, pomiar jest mało dokładny.

W drugim dniu wartości wynosiły 25-45 μg/m³. Rozkładały się na całym obszarze podobnie, najczęściej ok. 40-45 μg/m³. Wartości minimalne w centralnej części wykresu.

W trzecim dniu wartości wynosiły 26-36 µg/m³, najczęściej - ok. 26-28 µg/m³. Wartości maksymalne odnotowano na wschodzie i południu.







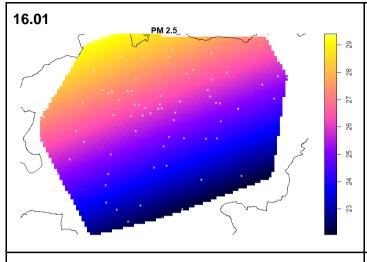
przesyłających brak informacji, pomiar jest mało dokładny.

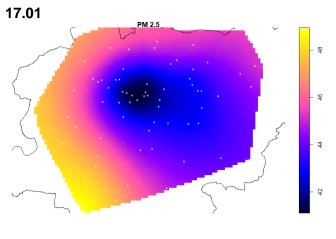
W drugim dniu wartości wynosiły 25-50 µg/m³.

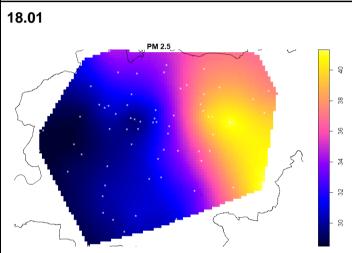
W pierwszym dniu wartości wynosiły 19-20 µg/m³. Maksymalne zanieczyszczenie odnotowano na wschodzie. Z powodu dużej liczby czujników

Rozkładały się na całym obszarze podobnie, najczęściej ok. 40-45 µg/m³. Wartości minimalne w centralnej części wykresu.

W trzecim dniu wartości wynosiły 30-50 μg/m³. Widzimy wyraźny podział na część wschodnią i zachodnią. Część wschodnia jest mniejsza, o wartościach ok. 40-50 μg/m³. Większa część zachodnia oscyluje między 30 a 35 μg/m³.





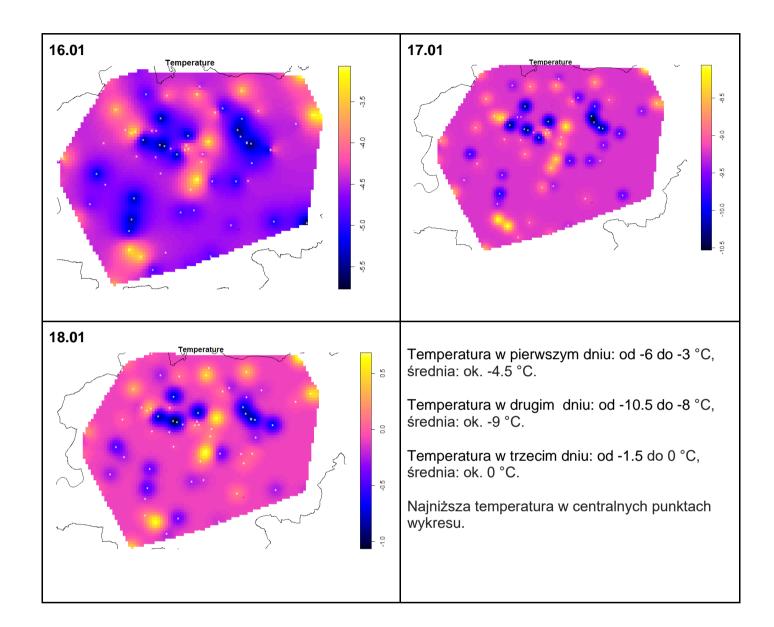


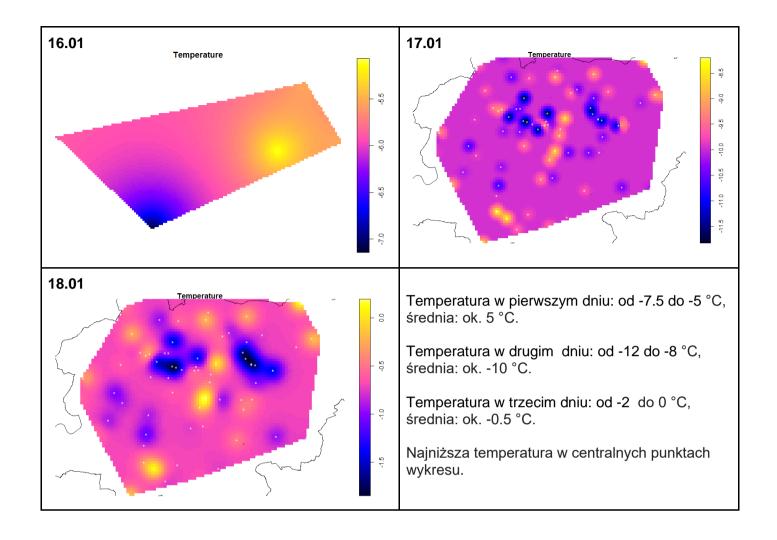
W pierwszym dniu wartości wynosiły 20-29 μg/m³. Wartości przedstawiały się poziomo i sekwencyjnie. Maksymalne wartości odnotowano na północy - ok. 27-29 μg/m³. Minimalne wartości na południu - ok. 20-25 μg/m³.

W drugim dniu wartości wynosiły 40-50 μg/m³. Wartości maksymalne na zachodzie, do 48-50 μg/m³. Wartości minimalne w centralnej i wschodniej części - ok. 42-44 μg/m³.

W trzecim dniu wartości wynosiły 28-45 μg/m³. Najniższe odnotowano na zachodzie - ok. 30-32 μg/m³, najwyższe na wschodzi, ok. 38-42 μg/m³.

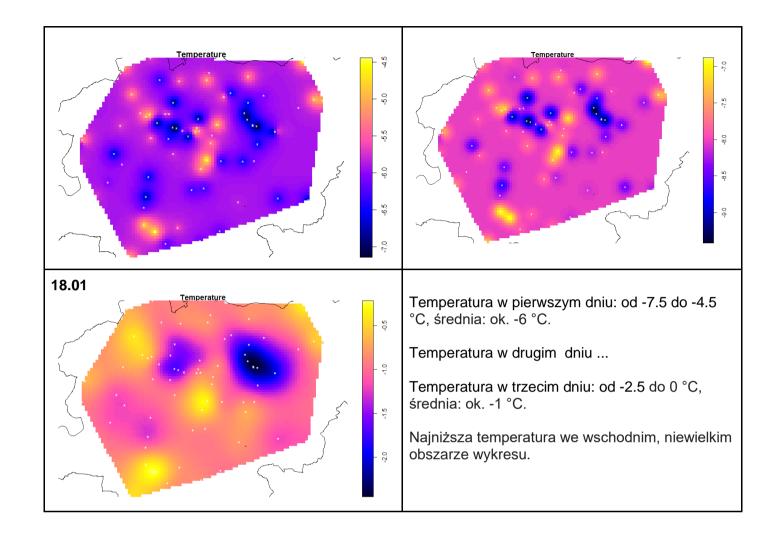
4.2. Temperatura



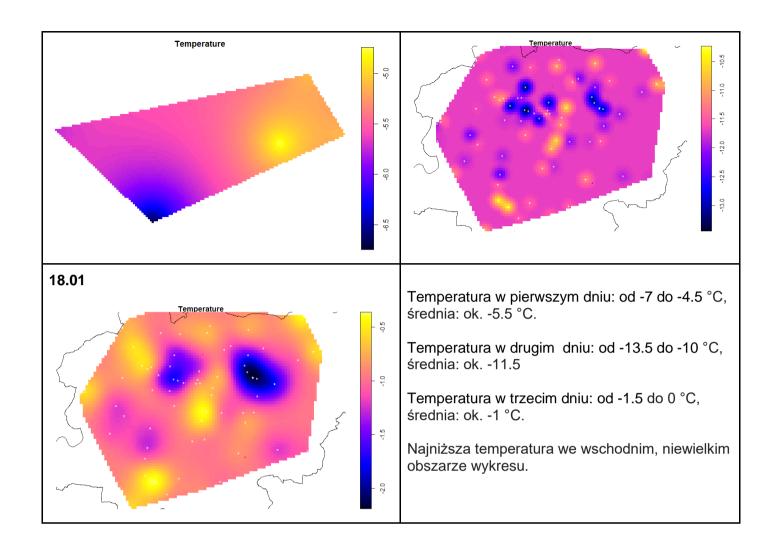


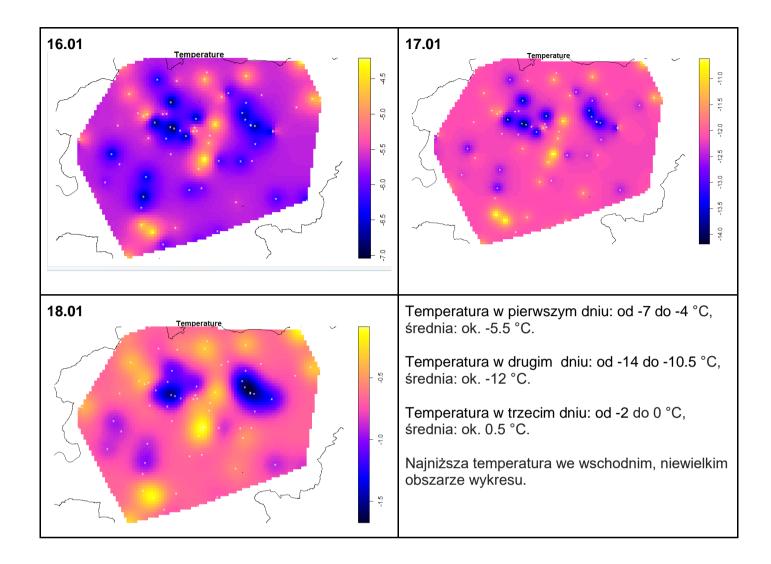
GODZ. 19:00

16.01 17.01

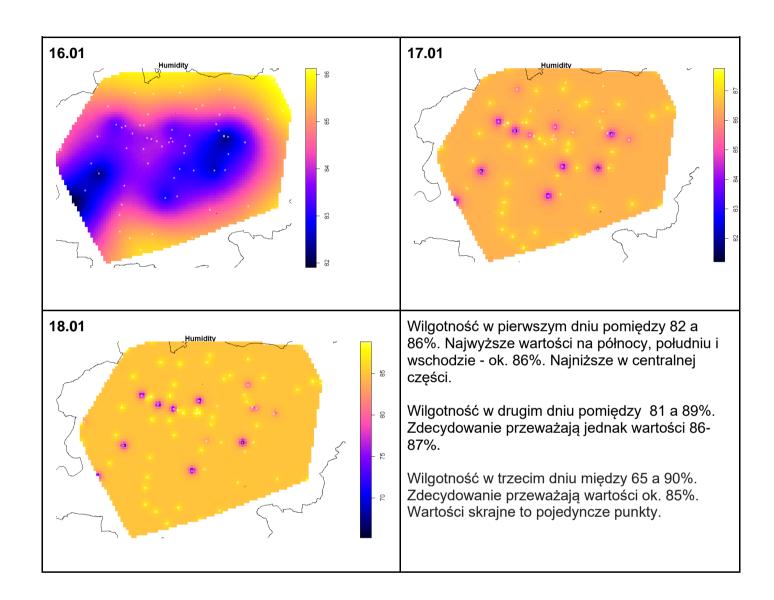


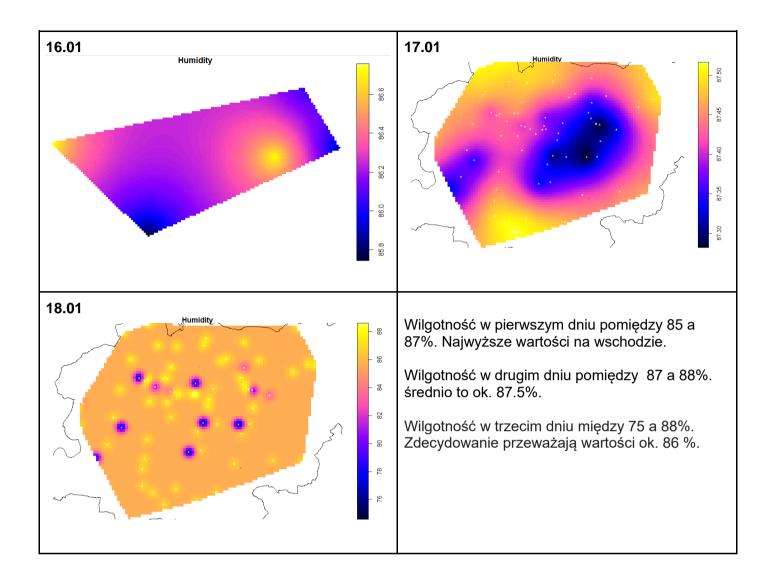
GODZ. 20:00

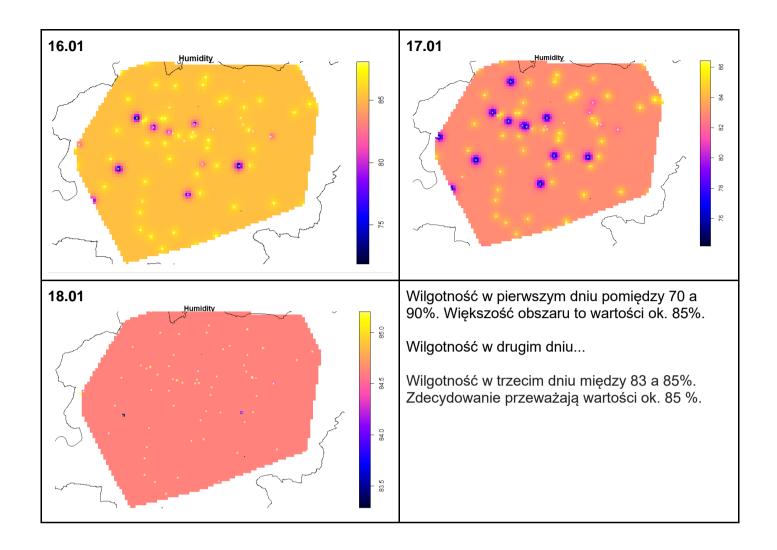


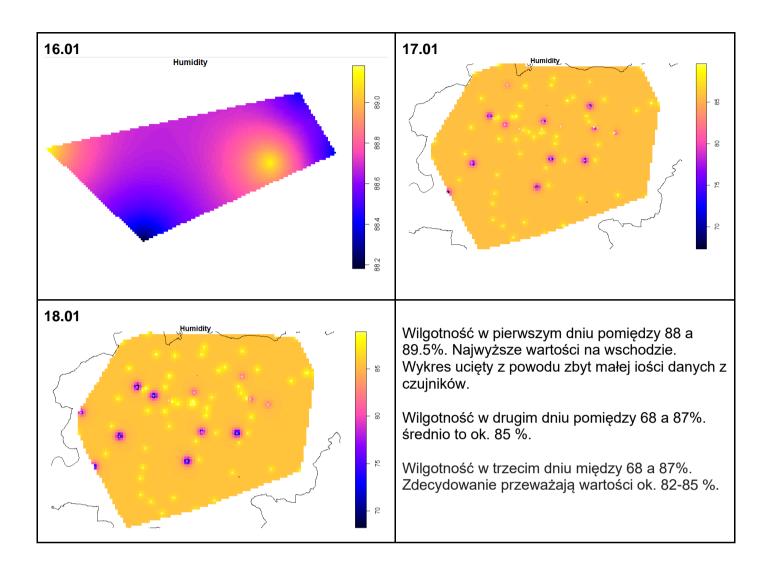


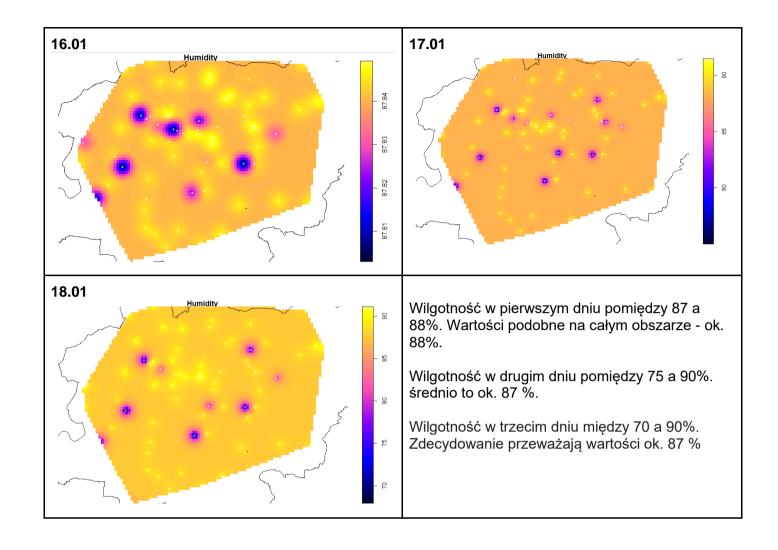
GODZ. 17:00











5. Wnioski

5.1. Godziny poranne i nocne

ZANIECZYSZCZENIA

Porównując godziny poranne i późno wieczorne (a więc takie, kiedy ulice zazwyczaj są puste), możemy zauważyć zdecydowaną przewagę wczesnej pory dnia. Widoczny jest niższy poziom zanieczyszczenia w powietrzu niż o północy. Prawdopodobnie o północy powietrze zawiera jeszcze dużo pyłów z całego dnia, z których przez noc się oczyszcza.

TEMPERATURA I WILGOTNOŚĆ

Najniższe wartości temperatury wystąpiły o godz. 6:00 - dochodziło nawet do -15 °C. O północy temperatura była zwykle nieco wyższa. Wilgotność przedstawia się podobnie na wszystkich pomiarach. Zwykle oscyluje w okolicy 80-90%. Najniższa występuje na zachodzie, natomiast najwyższa zwykle na północy.

5.2. Analiza dla biegaczy

ANALIZA ZANIECZYSZCZEŃ

Otrzymane dane skłaniają do kilku wniosków, przydatnych osobom aktywnym. Przede wszystkim należy bliżej przyjrzeć się zanieczyszczeniom.

W pierwszym dniu zanotowaliśmy duży wzrost wartości 2.5PM pomiędzy godz. 17:00 a 18:00. Mogło być to spowodowane dużym ruchem ulicznym i powrotem części osób z pracy do domu. W tym czasie nie odnotowano zmian temperatury, dlatego możemy wykluczyć związki powstałe z ogrzewania jako powód wzrostu. O godzinie 19:00 wartość wraca do liczby z godziny 17:00. Jest to ok. 20-30 μg/m³, co stanowi ok. 110-120% normy. Następnie wartość utrzymuje się na podobnym poziomie.

Sytuacja podobnie przedstawia się dla 17.01. Między godz. 17:00 a 21:00 wartość zanieczyszczenia powietrza utrzymuje się na podobnych poziomie. Najwyższa wartość (50 µg/m³, czyli 200% normy) występuje o godz. 19:00. Nailepiej sytuacja przedstawia się dla godziny 20:00.

W trzecim dniu, 19.01, mamy do czynienia ze stopniowym wzrostem zanieczyszczenia powietrza. Między godziną 17:00 a 20:00 cały czas stopniowo wzrasta do 50 μg/m³, by o godzinie 21:00 spaść.

WNIOSKI

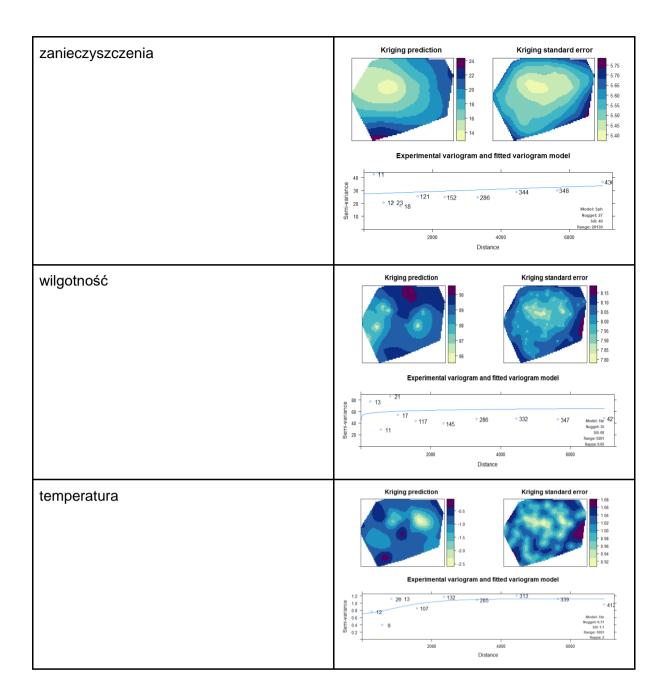
Analiza trzech dni to krótkie badanie, które pokazało różnorodne rezultaty, jednak możemy wysnuć wartościowe wnioski. Przede wszystkim zauważalne jest, że późniejsze godziny są lepszym terminem na aktywność fizyczną. We wszystkich trzech dniach zanieczyszczenie powietrza od godz. 20:00 miało umiarkowaną, bądź najniższą z badanych wartość. Warto także wziąć pod uwagę godzinę 21:00. W większości przypadków, im późniejsza godzina, tym czystsze powietrze.

Zdecydowanie najwyższe wartości występowały na wschodzie obszaru - w okolicach Nowej Huty. Zaskoczeniem jest, iż najniższe występujące wartości to punkty w centralnym Krakowie - w okolicach Krowodrzy i Śródmieścia. Kilka razy najwyższe wartości znajdowały się także na północy - Prądnik Biały. Badania wykazały, że zanieczyszczenie powietrza jest ściśle związane z temperaturą. Bardzo duży wzrost odnotowano w dzień, gdy temperatura mieściła się w granicach od -14 do -10 stopni. Wilgotność powietrza utrzymywała się na podobnych poziomach i nie odnotowano związku między nią a temperaturą lub zanieczyszczeniami powietrza.

6. Wykorzystane procedury

6.1. Analiza błędów

Oczywiście przed wyciągnięciem jakichkolwiek wniosków i przed dokonaniem głębszych analiz porównywałyśmy otrzymywane wykresy z ich odpowiednikami dotyczącymi błędów pomiarowych.
Przykładowe wyniki jakie otrzymałyśmy:



Możemy wywnioskować, iż poziom błędów w dużej części zależy od rozmieszczenia czujników. Tam, gdzie jest ich więcej, nasze pomiary są dokładniejsze oraz błędy mniejsze. Najbardziej wiarygodną więc częścią są dla nas okolice Starego Miasta, najmniej zaś wschodnie części Krakowa. Wykresy te jednak są dla nas ważną częścią, mówiącą o rzetelności i prawdziwości otrzymanych wyników.

6.2. Biblioteki

W projekcie wykorzystano liczne biblioteki z pakietu RStudio, które prezentują się następująco:

- rgdal
- maptools
- automap
- raster
- rgeos
- httr
- jsonlite

6.3. Komendy

1) Wczytanie danych z Airly

```
r<-
GET("https://airapi.airly.eu/v2/installations/neares
t?lat=50.0617022&lng=19.9373569&maxDistanceKM=10&max
Results=-1",add_headers(apikey = "KODAIRLY", Accept
= "application/json"))

jsonRespText<-content(r,as="text")
test10<-fromJSON(jsonRespText)
test15<-fromJSON(jsonRespText)
unique(test15$address$city)
test15_k<-test15[test15$address$city=="Krakow",]</pre>
```

2) Utworzenie zmiennych do przechowywania długości i szerokości geograficznej oraz wysokości i przypisanie ich z odpowiednich danych:

```
longitude<-test15$location$longitude
latitude<-test15$location$latitude
data15<-data.frame(longitude, latitude)
data15$elevation<-test15$elevation</pre>
```

3) Rysowanie mapek:

```
lat <- c(min(data15$latitude)-0.05,
max(data15$latitude)+0.05)
lon <- c(min(data15$longitude)-0.05,
max(data15$longitude)+0.05)
bbox <- make_bbox(lon,lat)
longitude<-test15_k$location$longitude
latitude<-test15_k$location$latitude
data15_k<-data.frame(longitude,latitude)
data15_k$elevation<-test15_k$elevation
b<- get_map(bbox,maptype="terrain",
source="osm",color="color",messaging = TRUE)
ggmap(b)+geom_point(data = data15,
aes(longitude,latitude)) + geom_point(data = data15_k, aes(longitude,latitude),col="Red")</pre>
```

4) Rysowanie konturów:

```
dzielnice<-st_read(dzielnice_Krakowa.shp)
dzielniceWGS84<-st_transform(dzielnice,crs = 4326)
plot(dzielniceWGS84)
(st_geometry(dzielniceWGS84))</pre>
```

5) Usuwanie granic wspólnych oraz przypisanie układu współrzędnych:

```
krakowWGS84<-st_union(dzielniceWGS84)
plot(krakowWGS84)
points (data15_k,pch=19)
points (data15,pch=1)

krakowUTM<-st_transform(krakowWGS84,CRS("+proj=utm
+zone=34N +datum=WGS84"))
data_spat<-
data.frame(lon=data15$longitude,lat=data15$latitude,
elev=data15$elevation)

coordinates(data_spat) <- ~lon+lat
proj4string(data_spat) <- CRS("+proj=longlat
+datum=WGS84")</pre>
```

6) Transformacja danych i testowanie:

```
data_UTM <- spTransform(data_spat, CRS("+proj=utm
+zone=34N +datum=WGS84"))

data15_ppp_e<-
ppp(x=data_UTM$lon,y=data_UTM$lat,marks=data_UTM$ele
v,window=as.owin(krakowUTM))

qt <- quadrat.test(data15_ppp)
plot(qt), qt
nn<-nndist(data15_ppp)
hist(nn)

ed<-density(data15_ppp)
plot(ed)
points(data15_ppp,pch="*",col="White")</pre>
```

7) Wykonanie mapy dla wysokości punktów pomiarowych

```
7.1) I sposób
```

Polega głównie na automatycznym krigingu dostępnych danych i wyrysowniu odpowiednich wartości.

```
data15_spdf<-
as.SpatialPointsDataFrame.ppp(data15_ppp_e)
elev_auto <- autoKrige(marks ~ 1, input_data =
data15_spdf) plot(elev_auto)
points(data15_ppp_e,pch="*",col="White")
plot(Window(data15_ppp_e),add=TRUE)</pre>
```

7.2) Il sposób

Utworzenie powyższej mapki, jednak z obramowaniem przyciętym do granic Krakowa. Do wycięcia owej siatki używamy obiektu spatial polygons, by później dla bardziej praktycznego użycia przekształcić ją na piksele.

```
bound<-st_as_sf(krakowUTM)
coord<-as.data.frame(st_coordinates(krakowUTM))
left_down<-c( min(coord$X), min(coord$Y))</pre>
```

```
right_up<-c( max(coord$X), max(coord$Y))</pre>
size<-c(100,100)
points<- (right_up-left_down)/size #wyliczenie</pre>
liczby punktÅłw
num_points<-ceiling(points) #zaokroÄ…glenie w gÅłrÄ™
eps<-10
num_points<-num_points+eps</pre>
qrid <- GridTopology(left_down, size,num_points)</pre>
gridpoints <- SpatialPoints(grid, proj4string =</pre>
CRS("+proj=utm +zone=34N +datum=WGS84"))
cropped_gridpoints <- crop(gridpoints,bound)</pre>
plot(cropped_gridpoints,add=TRUE,col="Red")
plot(Window(data15_ppp_e),add=TRUE)
spgrid <- SpatialPixels(cropped_gridpoints)</pre>
coordnames(spgrid) <- c("x", "y")</pre>
plot(spgrid)
plot(Window(data15_ppp_e), add=TRUE)
elev_auto <- autoKrige(marks ~ 1, input_data =
data15_spdf,new_data=spgrid,model = "Mat")
plot(elev_auto$krige_output[1], main='PM 2.5')
points(data15_ppp_e, pch="*", col="White")
plot(Window(data15_ppp_e), add=TRUE, border="White")
plot(elev_auto$krige_output[3])
points(data15_ppp_e, pch="*", col="White")
plot(Window(data15_ppp_e),add=TRUE,border="White")
```

8) Wykonanie mapy dla zanieczyszczenia powietrza pyłami 2.5 Pobieramy aktualne dane z czujników, a następnie wyciągamy interesujące nas dane (tu zanieczyszczenia) poprzez zmienną logic. Ważnym elementem jest też pozbycie się wartości nieznanych (NA) w celu korzystania z funkcji autoKrige.

```
current<-rep(NA,n_id)

for (i in seq(1,n_id)) {
   print(i)
   logic<-list_inst2[[i]]$current$values$name=="PM25"
   if (sum(logic)==1)
      current[i]<-
list_inst2[[i]]$current$values[logic,2]
}</pre>
```

```
data15_spdf<-
as.SpatialPointsDataFrame.ppp(data15_ppp_id)
data15_spdf$current<-current
miss <- is.na(data15_spdf$current)
pm25_auto <- autoKrige(current ~ 1, input_data =
data15_spdf[!miss,])
plot(pm25_auto$krige_output[1],main="PM 2.5")
points(data15_ppp_id[!miss,],pch="*",col="White")
plot(Window(data15_ppp_e),add=TRUE)
plot(pm25_auto)</pre>
```

9) Wykonanie mapy dla wilgotności powietrza Podobnie jak powyżej. Zmieniona jest tylko linijka w pierwszej pętli for na:

```
logic<-
list_inst2[[i]]$current$values$name=="HUMIDITY"</pre>
```

10) Wykonanie mapy dla temperatury powietrza Podobnie jak powyżej. Zmieniona jest tylko linijka na: logic<list_inst2[[i]]\$current\$values\$name=="TEMPERATURE"