PAPERS/THEORY SUMMARY

General idea:

- we have n distinct locations $s_1, \dots, s_n$ where $s_i = (\text{lat}_i, \text{long}_i)$, i.e. the stations
- there we record $Y_i(t)$ and $X_i(t)$, the PM10 and covariates
- we want to cluster them, so we define

$$\rho_t = \{S_1^{(t)}, \dots, S_{k_t}^{(t)}\} \quad (\Leftrightarrow) \quad c_i(t) = h \text{ if unit } i \text{ is assigned to cluster } h \text{ (at time } t)$$

(⇔) equivalent formulation

To solve/deal with this time dependence we can do basically in two ways:

(1) use models which incorporate time
(2) run a spatial model for each of the time instants

For (2) seems easy: the ppmSuite package does this (2) with also using covariates, so implementing (2) should be easy, like loop over time $t = 1 \to T$ and build a model at each instant. [illegible] we should use salso package maybe in the loop to estimate the best clusters. But we will see better when we study the packages.

models here on package ppmSuite:
- curve-ppmx
- gaussian-ppmx
- sppm

⇒ maybe we can split there ...

For (4) it was the main focus of paper 4 (Rossi and Quintana), so also there we should be ok as there was their package drpm. But that was not a published package, so there was no real/clear documentation (it's just the code in the folder papers/supplementary-material).
But it seems a small package (like 4 functions, 2 models) so should still be easy.

models from drpm package: — drpm-fit / maybe some others } ... and their/these

- 4 (3+4) functions for the models
- data finalization

but surely we need the final dataset → another task
(divided by weeks, water no NAs, ecc, standardize covariates)

TASK SPLIT

Maybe we can divide the functions/models to study, test, and start to experiment in R.
Also, on the ppm suite there are some papers associated to those functions, so we should study also them.
Well, who was assigned to that function studies it. In case of theory doubts ask Federico :)

For the models section the idea of the task is to understand
- which kind of input they want
- which kind of output they produce

For the data simulation task we just need to
- fill the NA gaps (or not?)
- scatter variables (PM10 and covariates) by week, as our time interval requests

Models from ppmm suite:

---

curve_ppmx YES MODEL 1 — *Gaussian PPMx Model for Functional Realizations*

models our data are functional but we could try also this

Description

curve_ppmx is the main function used to fit Functional Gaussian PPMx model.

Usage

see the example in the next page, or experiment with the command, to understand how y, z, ecc should be in our data case

```
curve_ppmx(y, z, subject,
           Xcon=NULL,Xcat=NULL,
           Xconp=NULL,Xcatp=NULL,
           PPM, M,
           q=3, rw_order=1, balanced=1,
           nknots,npredobs,
           Aparm, modelPriors,
           similarity_function=1,
           consim, calibrate,
           simParms,
           mh=c(1,1),
           draws=1100,burn=100,thin=1)
```

---

sppm YES MODEL 3 — *Function that fits spatial product partition model with Gaussian likelihood*

Description

sppm is the main function used to fit model with Guassian likelihood and spatial PPM as prior on partitions.

Usage

```
sppm(y,s,
     s.pred=NULL,
     cohesion,
     M=1,
     modelPriors=c(0, 100^2, 10, 10),
     cParms=c(1, 1.5, 0, 1, 2, 2),
     mh=c(0.5, 0.5),
     draws=1100,burn=100,thin=1)
```

---

gaussian_ppmx YES MODEL 2 — *Function that fits Gaussian PPMx model*

Description

gaussian_ppmx is the main function used to fit Gaussian PPMx model.

Usage

```
gaussian_ppmx(y, X=NULL, Xpred=NULL,
              meanModel=1,
              cohesion=1,
              M=1,
              PPM = FALSE,
              similarity_function=1,
              consim=1,
              calibrate=0,
              simParms=c(0.0, 1.0, 0.1, 1.0, 2.0, 0.1, 1),
              modelPriors=c(0, 100^2, 1, 1),
              mh=c(0.5, 0.5),
              draws=1100,burn=100,thin=1,
              verbose=FALSE)
```

they all have a similar structure so should be easy to study them once we understand one of them