General idea: we have n distinct locations $z_1, \dots, z_n$
($z_i = $ (lat, long) eg)

paper 3 pag 5

the goal is to define a model for the partitioning of z_i into $k_n(t)$ groups

so we define $\rho_n = \{S_1, \dots, S_{k_n(t)}\}$ the clusters identified, ie $i \in S_a \iff$ location z_i is in cluster a.
Or equivalently let $c_1, \dots, c_n$ with $c_i = a$ iff etc.

so maybe for us it will become $C(S_a, t)$?

Then the prior for ρ_n is

$$\pi_{\rho_n}(\rho) = P(\rho = \rho_n) \propto \prod_{a=1}^{k_n} C(S_a, z_a^*)$$

S_a will be like $\{1, \dots, n\}$ subset

this is a cohesion function, measures how likely those selected values of S_a would be clustered together a priori

this makes the clustering process location dependent

the set of locations that are in S_a, like if $S_a = \{1, 3, 9\}$ then $z_a^* = \{z_1, z_3, z_9\}$

we call this model/law as SPPM (spatial product partition model)

Then the idea is to set to each obs y_i for $i = 1 \to n$ a law/likelihood which is cluster dependent

paper 3 pag 16

$$y_i = Y(z_i) \mid \underline{\theta}, c_i \overset{ind}{\sim} f(\underline{\theta}^*_{c_i}) \quad \text{for } i = 1 \to n$$
$$\underline{\theta}^*_a \overset{iid}{\sim} G_0 \quad \text{for } a = 1 \to k_n$$

us actually have also "known" $\rho_n(t)$ as $k_n = |\rho_n(t)|$

$$\{c_i\}_{i=1}^n \sim \text{SPPM}$$

for example $\underline{\theta}^*_{c_i} = (\mu_{c_i}, \sigma^2_{c_i})$ and we set $f(\dots) = N(\dots)$

and we can include covariates with
let $y_i(t)$ the values observed (again for us)
let $x_i(t)$ the covariate set there
let $\beta^*_1, \dots, \beta^*_{k_n(t)}$ the cluster specific regression params

the $\theta^*_{c_i}$ of before

(wait, how can we have $k_n(t)$? a varying make the later inference ie, how will the work then? if the # k_n varies, not then the βs are tune as the same "maybe"

then we model
$$y_i = Y(z_i) \mid x_i(t), c_i(t), \beta^*, \sigma^2 \overset{ind}{\sim} N(x_i^T \beta^*_{c_i}, \sigma^2)$$

is all this temporal extension that makes me worry)

paper 3 pag 22

Here there are some more examples of models. but that is, note that the ways $\{c_i\}_{i=1}^n \sim$ SPPM, is equivalent to $\rho_n \sim$ SPPM, as from the ρ_n we can transform to get the c_i.

$\rho_n = \{S_1, \dots, S_{k_n}\}$
eg $\begin{pmatrix}2\\3\\7\end{pmatrix}\begin{pmatrix}1\\2\\4\\5\end{pmatrix}\begin{pmatrix}8\end{pmatrix}\begin{pmatrix}6\\9\end{pmatrix}$ $k_n = 4$

```
for i in 1:n
| c_i = which(i in S_j for j=1:k_n)
end
```

Ok, now, now we can understand better the model of our (in. thesis) most relevant paper, the paper 4.

paper 4 pag 9

Here they also provide the time extension! at each time

As they set $i = 1 \to m$ the units, ok, they model the partition as $\rho_t = \{S_1(t), \dots, S_{k_t}^{(t)}\}$ => NICE

or again $c_t = (c_1(t), \dots, c_m(t))$ as before.

So we want to study the P of getting a certain collection of partitions. So, assuming a MC memory) structure, we set our target to be as

$$p(\rho_1, \dots, \rho_T) = P(\rho_1(t) = \rho_1, \dots) = p(\rho_T \mid \rho_{T-1}) \cdot \dots \cdot p(\rho_2 \mid \rho_1) \cdot p(\rho_1)$$

markov prop

Then they use as prior for the partition structure this

$$\text{"}Pr(\rho \mid M)\text{"} = p_{\rho_m \mid M}(\rho \mid M) = P(\rho_m = \rho \mid M) = \underbrace{\frac{M^k}{\prod_{i=1}^{m} (M+i-1)}}_{\text{the consistent normalization const}} \prod_{w=1}^{k} \underbrace{(|S_w|-1)!}_{\text{the } \Gamma(|S|) \text{ told in the recent paper}}$$

We call this law CRP(M). But that law misses how we want to model the connection among ρ_{t-1} and ρ_t.

$\rho \sim CRP(M)$

So we introduce

$$\gamma_{it} = \begin{cases} 1 & \text{unit } i \text{ stays in the same cluster from } t-1 \text{ to } t \\ 0 & \text{otherwise} \end{cases}$$

we set then as random a $\gamma_{it} \overset{iid}{\sim} Ber(\alpha_t)$ — $\alpha_t = 0 : \rho_t \perp \rho_{t-1}$; $\alpha_t = 1 : \rho_t = \rho_{t-1}$

Then for better intuition, assuming we introduce $\gamma_t = (\gamma_1(t), \dots, \gamma_m(t))$ where $\gamma_i(t) \in \{0, 1\}$, to govern [illegible] which values we get from the $Ber(\alpha_t)$ for the γ_{it} we introduced above.

We call this model $tRPM(\alpha, M)$, and that is what eventually we will end in the model, about c_t, i.e. equivalently on the ρ_t.

$$
\begin{aligned}
Y_{it} \mid \boldsymbol{\mu}_t^\star, \boldsymbol{\sigma}_t^{2\star}, \boldsymbol{c}_t &\overset{ind}{\sim} N(\mu_{c_{it}t}^\star, \sigma_{c_{it}t}^{2\star}),\ i = 1, \dots, m \text{ and } t = 1, \dots, T, \\
(\mu_{jt}^\star, \sigma_{jt}^\star) \mid \theta_t, \tau_t^2 &\overset{ind}{\sim} N(\theta_t, \tau_t^2) \times UN(0, A_\sigma),\ j = 1, \dots, k_t, \\
(\theta_t, \tau_t) &\overset{iid}{\sim} N(\phi_0, \lambda^2) \times UN(0, A_\tau),\ t = 1, \dots, T, \\
(\phi_0, \lambda) &\sim N(m_0, s_0^2) \times UN(0, A_\lambda), \\
\{\boldsymbol{c}_t, \dots, \boldsymbol{c}_T\} &\sim tRPM(\boldsymbol{\alpha}, M), \text{ with } \alpha_t \overset{iid}{\sim} Beta(a_\alpha, b_\alpha),
\end{aligned}
\tag{5}
$$