RAPPORT DE STAGE

17 Mars 2021 – 22 Septembre 2021

Optimisation de l'effort d'échantillonnage dans le temps et dans l'espace

Etudiant

Noudéhouénou Freedich Madjid HOUESSOU

Superviseurs

Frédéric GOSSELIN, INRAE Nogent-Sur-Vernisson

Romain JULLIARD, CESCO – MNHN

Fabien LAROCHE, INRAE Toulouse

Antoine LEVEQUE, UMS PatriNat

Date de soutenance

27 Septembre 2021

Contents

Résumé

Les populations d'espèces sauvages ont des structures dynamiques. Elles sont largement distribuées. Le suivi de leurs tendances nécessite un échantillonnage. Le présent travail aborde la question de la répartition optimale de l'effort d'échantillonnage au fil du temps et dans l'espace afin de minimiser l'imprécision de l'estimateur de la tendance. Concrètement, nous partons des travaux de réalisés par Rhodes et Jonsen en 2011 dont nous relevons les limites et proposons une extension nécessaire pour l'application dans de réels écosystèmes gérés. Nous introduisons dans le plan d'échantillonnage, les notions de placettes non-permanentes et de strates. Le modèle avec placettes non-permanentes fait l'objet de simulation numérique. Les résultats obtenus traduisent un intérêt quantativement modéré de recourir aux placettes non permanentes dans le cadre d'un programme de suivi de la biodiversité. Nous discutons les perspectives ouvertes par ce travail en rapport à l'échantillonnage spatio-temporel dans les programmes de surveillance de la biodiversité.

Mots-clés :
modèle de croissance de Gompertz, distribution stationnaire, plan d'échantillonnage optimal, auto-corrélation spatiale, auto-corrélation temporelle, modèle auto-régressif.

Abstract

Species have dynamic structures. They are broadly distributed. Monitoring their trends calls for sampling. This work addresses how to allocate sampling effort over time and space in order to minimize the imprecision of the trend estimator. Concretely, we start from the work of Rhodes and Jonsen for which we identify the limitations and propose a necessary extension for the application in real managed ecosystems. We introduce the notions of non-permanent plots and strata into the sampling design. We compute the model with non-permanent plots. The results obtained reflect a quantitatively moderate interest of using non-permanent plots as part of a biodiversity monitoring program. We discuss the perspectives of this work related to spatio-temporal sampling in biodiversity monitoring programs.

Key-words:
Gompertz growth model, stationary distribution, optimal design, spatial correlation, temporal correlation, auto-regressif process.

Remerciements

D'abord, j'exprime mes remerciements et ma gratitude à mes encadrants Fréderic GOSSELIN, Romain JULLIARD, Fabien LAROCHE et Antoine LEVEQUE pour la confiance, le suivi, l'explication du contexte du stage, les conseils, les relectures et les réunions oh combien intéressantes les unes après les autres. Cette immersion dans le monde de la recherche appliquée via ce projet de stage a été très formateur pour moi. Mes remerciements particuliers à Frédéric et Fabien qui m'ont guidé et dont les apports et orientations m'ont permis de progresser dans mes travaux. Un merci à Frédéric pour sa disponibilité même certains week-ends. Malgré le télétravail conformément aux mesures liées à la COVID-19, vous avez toujours été présents pour m'accompagner.

J'adresse tous mes remerciements à l'endroit de tout le personnel du Muséum National d'Histoire Naturelle (Isabelle WITTE, Jean-Marc ALLART, Constance VIGNES pour ne citer que ceux-la) et de l'Université de Pau et des Pays de l'Adour ainsi qu'à la plateforme GENOTOUL de l'INRAE-Toulouse qui ont facilité ce stage.

Mes remerciements vont à l'endroit de Christian PAROISSIN, mon tuteur académique.

Je remercie tout le personnel de l'Université de Pau et des Pays de l'Adour (Marielle PLUMET, Marina HILD et autres) et de l'Université d'Abomey-Calavi au Bénin (CIPMA-UAC) qui ont permis et rendu facile cette mobilité.

Enfin, merci aux collègues du master MSID de l'Université de Pau et des Pays de l'Adour et aux amis de la résidence pour l'accueil et les moments partagés ensemble.

1. Intoduction

The Habitat Fauna Flora Directive in its article 11 requires each country to monitor species and habitats of community interest. As a result, it seems necessary for each country to set up a monitoring system for both terrestrial and aquatic biodiversity (habitats,vegetation, animals). In France, the *PASSIFOR* (Proposals for the Improvement of the FOREST Biodiversity Monitoring System) and the "Terrestrial Biodiversity Monitoring Project" supported by the French Biodiversity Office have been designed for this purpose. The former aims at assessing the possibility of designing an effective system for monitoring forest biodiversity at different scales (national, regional, local) and to draw the practical modalities according to one or more scenarios (http://www.gip-ecofor.org/projet-passifor/). The latter is *"designed as an operational device, intended to answer the questions posed by public policies and society, to provide reliable and up-to-date information to monitor the state of biodiversity in connection with environmental pressures and that contributes to the evaluation of the conservation efforts that are made"* (Lévêque pers. comm., UMS PatriNat online).

Biodiversity monitoring refers to a system of regular observations of ecosystem over time, informing about biodiversity state with the purpose to detect and assess the trends of quantities like species richness, species diversity, species abundances (Mitusova 2006, Gosselin et al. 2007). It has an important role to play in managing environment and understanding population dynamics (Lindenmayer and Likens 2010). Species are made of individuals forming populations that have dynamic structures. Individuals are usually very difficult to count and are broadly distributed. Thus, the monitoring of their trends may cost enormously in time and financial resources (Nielsen et al. 2009). So ecological monitoring calls for sampling in order to estimate the biological quantities of interest. Moreover, monitoring usually consists in measuring environmental parameters and sampling species in small, pre-defined locations, called *plots*. There are several strategies to tackle the sampling in monitoring. In particular, the plots can be identical for each measurement campaign (permanent plots) or be may change from one campaign to the other. The most commonly used method is the first one (Urquhart and Kincaid 1999). In addition, the effort can be spread over many plots visited less often or over few plots visited frequently. The optimal choice between these last two alternatives depends on the spatial and temporal correlation of what is measured (Rhodes and Jonzén 2011). Rhodes and Jonsen research used a simplified spatio-temporal model to explore how to allocate the sampling effort among spatial and temporal replicates to minimize uncertainty in trend estimates. They showed that *"allocating sampling effort among spatial and temporal replicates depends on the spatial and temporal correlations in population dynamics and environmental variation"* (Rhodes and Jonzén 2011). Particularly, when spatial correlation is low and temporal correlation is high, the best option is to sample many sites infrequently, particularly when observation error and/or spatial variation in temporal trends are high. And when spatial correlation is high and temporal correlation is low, the best option is likely to be to sample few sites frequently, particularly when observation error and/or spatial variation in temporal trends are low.
The aim of this research is to determine the optimal distribution of the sampling effort in time and space using an extended version of Rhodes and Jonzen's framework. Concretely, we will emphasize the introduction of non-permanent plots and introduction of strata. Then we will also clarify the limits and simplifying assumption of the original Rhodes and Jonzen model and propose necessary extensions for application in real managed ecosystems.

2. Definitions of some concepts

2.1. Design methods

For finite population sampling there are two main modes of inferences which are *design-based* and *model-based* (Little 2004). *Design-based inference* estimates characteristics of a population only from the probabilistic nature of the design or sampling plan with asymptotic properties. It gives a good estimators in larges samples but is limited for small samples adjustment. It is not based on a data-based model. While *model-based inference* is based on a data-based model. If the model corresponds well to the model that generated the data, it gives an estimators with good properties based on likelihood or Bayesian principles (Little 2004). In addition, the *design-based approach* is often used in classical survey sampling, whereas the *model-based*

approach is used in geostatistics and in time series analysis in research (Gore 2008).
In this work, we used *model-based method* as Rhodes and Jonsen because it allows to take account of spatial and temporal structures of the data. Indeed, *in the model-based approach, the weights of the data are determined by the covariances between the observations, which are given by the model as a function of the coordinates of the sampling locations* (de Gruijter et al. 2006).

2.2. Spatial correlation vs temporal correlation

The *spatial correlation* reflects the existence of a resemblance relationship between the information collected on two close sites. While the *temporal correlation* reflects the resemblance between the information collected at two different dates on the same site.

2.3. Asymptotic state vs stationary state

The *asymptotic behavior* is the behavior of the process when t tends to infty. Whereas *stationary behavior* is the one where the mean and the variance do not dependent on t, i.e the dynamics has been running for a long time before the beginning of the survey. Here, we used *asymptotic behavior* for probabilistic model and *stationary behavior* for statistic model.
In this research, both of them were the same. Indeed,we parametrized the behavior of initial population from start in the statistic model under stationarity assumption as the same as the asymptotic behavior of the probabilistic model.

3. Materials and methods

This chapter will be subdivided into three parts. First, we will present Gompertz model which serves as a basis to both Rhodes and Jonzen's and the present work. Then, we will present the monitoring strategy used which is different from Rhodes and Jonzens's and finally the experiment.

3.1. Gompertz state-space (GSS) model : definition and notations

The Gompertz State-Space (GSS) model is a stochastic process used to depict time-series of population abundances. It combines density-dependence and environmental process noise (environmental variation). The population dynamics are represented by the discrete-time stochastic Gompertz model like :

$$N_{i,t} = N_{i,t-1}exp(-0.5\sigma^2 + \gamma(lnK_{i,t-1} - lnN_{i,t-1}) + u_{i,t}) \tag{1}$$

with $N_{i,t}$ the positive, real-valued represented the abundance of sub-population i at time t; $K_{i,t}$ the equilibrium abundance of sub-population i at time t; γ the strength of density-dependence; $u_{i,t}$ the stochastic environmental variation in the population growth rate for sub-population i at time t, described by a S-dimensional gaussian vector with mean zero and variance-covariance matrix $\mathbf{\Sigma}$. We defined $\mathbf{\Sigma}$ as Rhodes and Jonsen such that environmental variation among sub-population was spatially correlated, with $Var(u_{i,t}) = \sigma^2$ and $Cov(u_{i,t}, u_{j,t}) = \rho^{d_{i,j}}\sigma^2$. $-0.5\sigma^2$ makes sure that the expected growth rate was the same as the growth rate for the equivalent deterministic version of the model (Rhodes and Jonzén 2011).

On a logarithmic scale, the Gompertz model becomes

$$lnN_{i,t} = lnN_{i,t-1} - 0.5\sigma^2 - \gamma(lnN_{i,t-1} - lnK_{i,t-1}) + u_{i,t} \tag{2}$$

Rhodes and Jonsen define the quantity $\epsilon_{i,t} = lnN_{i,t} - lnK_{i,t}$ which we call here deviations from local equilibrium (DFLE).
If $\gamma = 0$, we have $lnN_{i,t} = lnN_{i,t-1} - 0.5\sigma^2 + u_{i,t}$. It represents the case of *density-independence*. Whereas if γ=1, we have $lnN_{i,t} = lnK_{i,t-1} - 0.5\sigma^2 + u_{i,t}$. It means that there is not temporal correlation. In this case, the population abundance $lnN_{i,t}$ fluctuates around carrying capacity $lnK_{i,t-1}$. And $\gamma > 1$ is for species with *chaotic behavior*.

3.2. Monitoring strategy and scenarios

Sampling method is one of the key points in ecological monitoring (de Gruijter et al. 2006 and Vos et al. 2000). The sampling method used by Rhodes and Jonsen in their work was to randomly select S sub-populations among the population P and to collect abundance at regular intervals in these same locations. That means that they used only *pure panels* (de Gruijter et al. 2006). One consequence is that there is not one-shot plots. Here, we choose to adopt *rotational sampling* which allow *revisited designs*. In other words, this strategy allows us to have both permanent and non-permanent plots. Concretely, permanent plots are sampled at regular sampling time whereas non-permanent plots are sampled randomly. In this work, we choose to sample non-permanents plots only one time. A potential advantage of this strategy is its flexibility and better spatial coverage (Gore 2008). Note that there is different *revisited designs* and *sampling methods*. People can get more information about them in de Gruijter et al. (2006) and McDonald (2003).

	Times						
Plots	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Permanents Plots							
1	X			X			X
2	X			X			X
3	X			X			X
4	X			X			X
5	X			X			X
Non Permanent Plots							
1	X						
2							
3	X						
4							X
5				X			
6							
7							
8				X			
9							X

Figure 1 : Illustration of sampling design synchronized permanent plots vs unsynchronized non-permanent plots

Rhodes and Jonsen in their paper considered different scenarios. The first ones in which they supposed that there is not observation error in the estimating of the abundance and variation in their temporal trends. The aim in this case was to identify the best monitoring program based on the parameters γ (temporal correlation and the strength of density-dependence), ρ (spatial correlation) and σ (environmental variation). And the second in which they supposed that there is an observation error and/or variation in the log-linear temporal trends. The best monitoring program in this case was identified as a function of γ, ρ, σ, σ_{obs} and σ_{trend} where the authors assumed that the abundances $lnN_{i,t}$ are observed with an error ($\sigma^2_{obs} > 0$).
In this work, we just considered two scenarios according to our monitoring strategy. A first one in which

there is only permanent plots and a second in which there is a mixture of permanent and non-permanent plots. The aim was to identify the best monitoring strategy depending on the parameters mentioned above and the proportion of non-permanent plots in the sample.

3.3. Experiments

For the sake of simplicity, we considered the simplify form of the Gompertz stochastic model. We assumed that there is no temporal variation in the log-linear trend equilibrium population sizes, r ($\sigma^2_{trend} = 0$). We also assumed that the distribution of $lnN_{i,t}$ is stationary which means that the dynamics has been running for a long time before the beginning of the survey. Then, like Rhodes and Jonsen, we can use the *generalised least squares (GLS)* formulas for estimating the variance of the log-linear temporal trend, r. In this case, the MVUE (minimum variance unbiased estimator) of r is the one we were looking for. Compared to the work of Rhodes and Jonsen , the innovation here is the introduction of non-permanent plots with one sampling time. Note that several combinations of μ, σ and γ are possible in the intercept of the model.

Specifically, let defined S the total number of sites, t_{max} the maximum monitoring time, S_p the total number of permanent sites, S_{np} the total number of non-permanent sites, T_p the number of visits per permanents sites and T_{np} the number of visits per non-permanent sites, T the average number of visits per sites and p the proportion of non-permanent sites. $B = S_p \times T_p + S_{np}$ is the cost of one scenario. The variance of the estimated temporal trend is then

$$Var(\hat{r}) = [(\mathbf{X}'\mathbf{\Phi}\mathbf{X})^{-1}]_{3,3} \tag{3}$$

where $\mathbf{X}$ is a design matrix of length $B \times 3$. $\mathbf{\Phi}$ is a $B \times B$ matrix specifying the variance-covariance of the residuals of the generalised least squares model, with $S = S_p + S_{np}$ and $T = (1-p)T_p + p$. The explicit form of the variance-covariance matrix $\mathbf{\Phi}$ is described in the appendix A (precisely eq.A35). Then we were able to identify optimal sampling designs (those which had the lowest sampling variance for estimating the log-linear temporal trend) and to identify how this varied with the parameters γ (temporal correlation), ρ (spatial correlation) and p (proportion of non-permanents plots). To do so, we considered a spatial domain with 10×10 square landscape (spatial unit). t_{max} were fixed at 30 years and B at 1500, constant. B and t_{max} in our work were bigger than those used by Rhodes and Jonsen because our monitoring is at national scale. We considered all combinations of $1-\gamma \in \{0.2, 0.4, 0.6, 0.8, 1\}$ and $\rho \in \{0, 0.2, 0.4, 0.6, 0.8\}$. Note that T and p have been set up in order to have the cost B. They were chosen in the space of the divisors of B. Thus, we had the following sampling designs : $T = 2, 3, 5, 6, 10, 15, 30$ and $p = 0, 0.5, 0.8, 0.9, 0.96$. The environmental variation and the variation of observation error were fixed ($\sigma^2 = 1$ and $\sigma^2_{obs} = 0.1$). We calculated the standard error $se(\hat{r})$ for precise values and γ, ρ, T and p and that for each crossing γ/ρ, we took the minimum T and p among these discrete values. Concretely, like Rhodes and Jonsen, we simulated, 100 times, the spatial location of the S sub-populations i surveyed by randomly located them on the landscape. For each replicate, we calculated the sampling variance of the log-linear temporal trend estimate. The expected sampling variance was then estimated as the mean of the sampling variance of the 100 replicates and we determined the optimal sampling design as those which has the minimum sampling variance of the trend estimate. (T_{min}, P_{min}) is a couple of parameters T and P corresponding to the optimal sampling design.
In order to facilitate decision-making, we introduced other indicators such as standard error $se(\hat{r})$ and the ratio of standard error between the case permanent plots and mixtures of permanent and non-permanent plots ($\Delta_{se} = \frac{se_{npp}(\hat{r})}{se_{pp}(\hat{r})}$) with $se_{pp}(\hat{r})$, the standard error for the case of permanent plots and $se_{npp}(\hat{r})$, the standard error for the case of non-permanent plots (see appendix C for resume of the different steps of the simulation).

4. Results

First, we analyzed the work of Rhodes and Jonsen in order to identify its limitations. Then, we extended them through the introduction of non-permanent plots and strata.

4.1. Analysis and identification of the limitations of Rhodes and Jonsen's work

The assumption of stationarity is obviously the first limit of the work of Rhodes and Jonsen. Indeed, in real life, monitoring does not necessarily start in a equilibrium state. The dynamics may have changed recently. Second, by fitting the stationary model, they assumed that γ (temporal correlation), ρ (spatial correlation) and σ (environmental variation) were perfectly known at their true value before, which is not true in real life in which these parameters have to be estimated. Moreover, Rhodes and Jonsen did not derive the mean structure of the deviations from local equilibrium. They also did not write the complete expression of the deviations from local equilibrium, $\epsilon_{i,t}$. They forgot some terms. Although these terms do not affect the variance-covariance matrix, they are contained in the structure of the mean and therefore should be taken into account in case of a complete fit to a data set. In addition, Rhodes and Jonsen fit a linear regression model $Y_t = f(t)$ without intercept where $Y_t = logN_{i,t} - logN_{i,0}$ (see eq.A37a and A54a for the expression of the model). We derived the variance-covariance matrix of this model in appendix A obtained a result different from the one used by Rhodes and Jonsen, which is another limit of their work.

4.2. Explicit temporal trajectories of the model

With the hypothesis and the methodological choice (stationarity and a type of regression $lnN_{i,t} - lnN_{i,0} = f(t)$) made by Rhodes and Jonsen in their work, they rule out the question of initialization at the beginning ($t = 0$) and only consider a simplified subspace of the possible temporal trajectories of $N_{i,t}$. Here we do not make these assumptions (see details about initialization of the process in section 4.3.2). Moreover, with the introduction of non-permanent plots, it is not possible to estimate the trend after only one visit using Rhodes and Jonsen model. Consequently, we opted for regression $Y_t = f(t)$ where $Y_t = logN_{i,t}$ with intercept. This regression model has the advantage to take account of the initial data and therefore would give a better estimators of other parameters of the model (environmental variation σ for example) especially when the dynamic is non-stationary. In this section, we focused on the right expression of the deviations from local equilibrium and their mean and variance-covariance structures which either did not appear or were potentially not true in Rhodes and Jonsen work.

Like Rhodes and Jonsen in their work, we assumed that our hypothetical population P is subdivided into several spatially distinct sub-populations i and that the dynamics of each of them follows a density-dependent stochastic Gompertz model (Rhodes and Jonzén 2011). $\mathbf{\Sigma}$ is defined such that the environmental variation between sub-populations is spatially correlated. The environmental variation of one sub-population i at time t is supposed to be unchanged ($Var(u_{i,t}) = \sigma^2$) and the covariance of two different sub-populations, i and j, at the same point in time, t, is equal to $\rho^{d_{i,j}}\sigma^2$ ($Cov(u_{i,t}, u_{j,t}) = \rho^{d_{i,j}}\sigma^2$), with ρ, the strength of spatial correlation in growth rates and $d_{i,j}$, the distance between two sub-populations i and j. ρ and γ are assumed to be independent and constant belonging to]0,1[. There is not interaction between spatial and temporal correlations. Environmental variation, $u_{i,t}$, are assumed to be independent with $N_{i,0}$ (the initial abundance of the sub-population i). The environmental variation at different time steps are assumed to be uncorrelated ($Cov(u_{i,t}, u_{i,s}) = 0$ with $t \neq s$).

In real life, population abundances are rarely known exactly and can be variable. In order to account for their effects on sub-population abundance estimates, observation errors and variation in the temporal trends among sub-population i are modeled by two normally distributed variables $v_{i,t}$ and η_i with means 0 and variance σ^2_{obs} and σ^2_{trend} respectively (see after appendix A) (Rhodes and Jonzén 2011). Observation errors are assumed to be independent for different sub-populations i and j or at different times t and $t - s$ for the same sub-population i. The same independence assumption has been made between ecological deviations from local equilibrium and observation errors. In addition, environmental variation and observation errors are assumed to not depend on variation in temporal trends of each sub-population.

Next, like Rhodes and Jonsen, we assumed that at logarithmic scale, the equilibrium populations have a deterministic temporal trend and a variable rate of change for each sub-population i. That means that $lnK_{i,t} = lnK_{i,t-1} + r_i = lnK_{i,t-1} + (r + \eta_i)$ where r is deterministic and η_i is an independent and normally distributed random variable with mean 0 and variance σ^2_{trend}. From eq.2, the recursion relationships for the deviations from local equilibrium for site i at time t is a linear autoregressive of order [AR(1) process] time series model defined as follows

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - \eta_i - 0.5\sigma^2 \tag{4}$$

$\epsilon_{i,0}$ where $i \in \{1, ..., S\}$ with S the total number of sites, is supposed to be a gaussian vector. So, $\epsilon_{i,t}$ with $i \in \{1, ..., S\}$ is a ST-dimensional gaussian vector, with T, the total number of visits, as a linear combination of gaussian variables. If there is not variation in the temporal trends among sub-population ($\sigma^2_{trend} = 0$), the rate of change of the equilibrium population sizes for each sub-population i is unchanged. In this case, the recursion relationships become

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - 0.5\sigma^2 \tag{5}$$

The derived mean of the deviations from local equilibrium (DFLE) is :

$$Mean(\epsilon_{i,t}) = (1-\gamma)^t Mean(\epsilon_{i,0}) - (r + 0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{6}$$

The derived asymptotic and non-asymptotic variance-covariance matrix of the deviations from local equilibrium is $\mathbf{\Pi}$ (see appendix A16 for the description of $\mathbf{\Pi}$).

4.3. Mathematical extension of Rhodes and Jonsen's model

4.3.1. Introduction of strata

In ecological monitoring, it is not uncommon to use strata in the sampling design to separate several types of statistical populations. For example, Carvalho et al. (2016) proposed a framework based on stratification to design optimized multispecies-targeted monitoring networks over large areas. Another part of our work consisted in extending the results obtained by Rhodes and Jonsen by introducing strata. Let assume that the population is divided into g strata with different trend parameters. Each sub-population i of each stratum is submitted to a different density-dependence, with γ_g, the strength (intensity) of that dependence in the strata. Each sub-population i of each stratum has a different trend r_i, and that these trends are normally and independently distributed with mean r_g and variance $\sigma^2_{trend,g}$. For the sake of simplicity, ρ and σ are supposed to be unchanged from one stratum to another. If not, it would be difficult to write the right expression of the variance between deviation from local equilibrium of the sub-population i at time $t-1$ and environmental variation of the sub-population j at time $t-s$, $Cov(\epsilon_{i,t-1}, u_{j,t-s})$. From eq.4 the deviations from local equilibrium becomes

$$\epsilon_{i,t} = (1-\gamma_{g(i)})\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2 \tag{7}$$

When temporal and spatial correlation are identical, eq.7 becomes

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2 \tag{8}$$

The analytic expressions of mean and variance-covariance are developed in appendix A.

4.3.2. Statistical space-time model for estimating the trend on abundance $N_{i,t}$

In order to evaluate potential dynamic designs, it is important to specify a space-time model with a low estimation error (Wikle and Royle 1999). This model must account for spatio-temporal variability and correlation in the environmental process (Wikle et al. 2019). For this reason, the prior part of our work consisted in taking again the calculations of Rhodes and Jonsen while relaxing the assumption of stationarity and writing the explicit form of the statistical model of the process. We assumed that the dynamic is not stationary. Our goal is to model the log-linear temporal trend in the population sizes over time. To do so, we gave a specific form to the initial states (initial population) such as initially the population is away from equilibrium and non-stationary. Then the structures of the initial deviations from local equilibrium, $\epsilon_{i,0}$, were parametrized as follows

$$Mean(\epsilon_{i,0}) = c \tag{9}$$

$$Var(\epsilon_{i,0}) = \frac{\sigma_0^2}{1-(1-\gamma)^2}(\sigma^2 + \sigma_{trend}^2(1 + \frac{2-2\gamma}{\gamma})) \tag{10}$$

$$Cov(\epsilon_{i,0}, \epsilon_{j,0}) = \frac{\sigma_0^2}{1-(1-\gamma)^2}(\rho^{d_{i,j}}\sigma^2 + \sigma_{trend}^2(1 + \frac{2-2\gamma}{\gamma})) \tag{11}$$

and

$$Cov(\epsilon_{i,0}, \eta_i) = \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2 + \sigma_{trend}^2(1 + \frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} \tag{12}$$

with c a constant and σ_0 positive. The initial carrying capacity $lnK_{i,0}$ were supposed to be deterministic and defined as follows

$$lnK_{i,0} = \mu + \omega x_i \tag{13}$$

The conditional form of the regression model is :

$$lnN_{i,t} = \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + (1-\gamma)^t(c + \frac{r + 0.5\sigma^2}{\gamma}) + \eta_i t + \epsilon'_{i,t} \tag{14}$$

where $\epsilon'_{i,t}$ is a centered multivariate gaussian variable defined as the difference between deviations from local equilibrium and its average value. The term $\eta_i t$ is deterministic and will be contained in the mean of $lnN_{i,t}$. Conversely, in the marginal version, the term $\eta_i t$ is stochastic and then is be contained in the residuals. This model can be written as follows:

$$lnN_{i,t} = \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + (1-\gamma)^t(c + \frac{r + 0.5\sigma^2}{\gamma}) + E_{i,t} \tag{15}$$

The residual here is $E_{i,t}$ and is defined as follows : $E_{i,t} = \epsilon'_{i,t} + \eta_i t$. The parameters that will have to be estimated in both two models are r, γ, ρ, σ, μ, ω, c, θ, σ_{trend} and σ_0 where c, θ and σ_0 are the initialization parameters (see Table 1).

Table 1 : List of parameters

Parameters	Definition
r	log-linear temporal trend
γ	Temporal correlation
ρ	Spatial correlation
σ	Environmental variation
μ	intercept of the initial carrying capacity
ω	slope of the initial carrying capacity
σ_{trend}	variation in log-linear temporal trend
σ_0	initialization parameter
c	initialization parameter
θ	initialization parameter

Due to the dependencies and variability in environmental process, the statistic model include a nonlinear structure in the mean, a precise parametrized form of variance-covariance, or even random effects whose parameters will have to be estimated under the non-stationary assumption. In appendix A, we derived the means and variance-covariance matrices of each model (with and without observation errors). The variance-covariance matrix of the residuals of the conditional model is identical to the variance-covariance matrix of the residuals of the statistical model without site effects (variation in the temporal trends among sub-populations). The means and variance-covariance of the models share same parameters, which makes them difficult to estimate with classical statistical tools. For estimating of the parameters of these models, some use methods such as Laplace approximation, quasi-likelihood, generalized estimating equations, pseudo-likelihood, and penalized quasi-likelihood through TMB (Template Model Builder) whereas other use Bayesian framework (Wikle et al.(2019)). Furthermore, due to complexity problems, it is desirable to consider both marginal and conditional forms of the model and estimate their parameters because one could be faster than the other.

Under stationary assumption, $\epsilon_{i,0}$ has the asymptotic behavior of the probabilistic model. That means that if $\sigma_0 = 1$, $c = \frac{-(r+0.5\sigma^2)}{\gamma}$ and $\theta = -\frac{\sigma_{trend}\sqrt{1-(1-\gamma)^2}}{\gamma\sigma_0\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}$, we are in the stationary regime and the regression model is a type of linear mixed-effects model. Consequently, the conditional regression model becomes

$$lnN_{i,t} = \mu + \omega x_i - \frac{r+0.5\sigma^2}{\gamma} + rt + \eta_i t + \epsilon'_{i,t} \tag{16}$$

Whereas the marginal model becomes

$$lnN_{i,t} = \mu + \omega x_i - \frac{r+0.5\sigma^2}{\gamma} + rt + E_{i,t} \tag{17}$$

Based on eq.17, the first column of the matrix $\mathbf{X}$ contained only 1 (for intercept). The second one contained the values of the variable x_i in each site at each time steps. The third column was for the variable time t. For, numerical implementations, we focused on marginal form of the model following Rhodes and Jonzen approach. Note that there are several methods to estimate parameters of this kind of model. We will not develop them here. Ones use generalised linear squares (GLS) for estimating parameters for the stationary form of the model because GLS explicitly accounts for the dependence in the errors (Wikle et al.2019). Dennis et al.(2006) have developed *maximum likelihood* and *restricted maximum likelihood* approaches to parameter estimation for this Gompertz stochastic model under stationary assumption. Also Wikle and Royle (1999), Rao and Toutenburg (1995) and Bates and Pinhero (2000) have correctly addressed the problem of estimating parameters in linears models, KK et al. (1998) and Faraway (2006) in non linears models.

4.4. Computation of non-stationary model

We first tried to compute the non-stationary model and estimate the parameters through *TMB* (Template Model Builder). But, we faced an error while simulating multinormal distribution of $lnN_{i,t}$. Which led us to Bayesian modeling. On *greta*, the MCMC estimators obtained after $N = 1000$ iteractions were not convergent. For example, the figure 2 show the convergence of the parameter $logit(\rho)$. The graphs did not mix well. Moreover, we got 23% of bad behaviour of the hamiltonian used in greta in the sampling phase, which indicates we cannot trust the output blindly. *Nimble* might give a result. Due the relatively short time that remained and the impossibility of being able to compare these results with those of the previous methods, we decided to implement the stationary model using Rhodes et Jonsen approach (see details of the simulation in section 4.5).

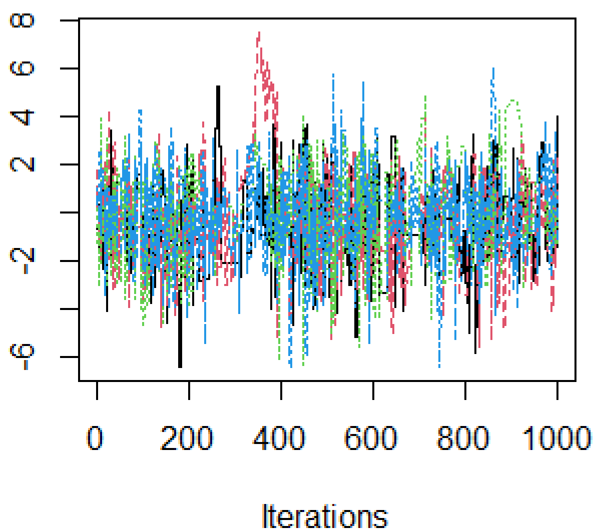

Fig 2 : convergent curve of $logit(\rho)$ with greta for N=1000 iteractions. S=4 and T=4

4.5. Computation of stationary model : results of experiment

We describe the scenario where there were only permanent plots. When spatial correlation was high (ρ high) and temporal correlation low ($1-\gamma$ low), the optimal sampling times was $T_{min} = 30$ (the maximum possible value). That means that allocating survey effort to temporal replicates was preferentially better (fig.3). While, when spatial correlation was low (ρ low) and temporal correlation high ($1-\gamma$ high), the optimal sampling times was $T_{min} = 2$ (the mminimum possible value). Surveying in two times was the best strategy. That means that it would be better to allocate survey effort to spatial replicates (fig.3). Moreover, when spatial correlation was high (ρ high) and temporal correlation high ($1-\gamma$ high), surveying in ten times out of thirty ($T_{min} = 10$) was the best strategy (fig.3). The standard deviation of the estimated trend for the best strategy was lowest when spatial and temporal correlations were low and highest when spatial and temporal correlations were high (fig.3). Furthermore, the preferred strategy did depend on spatial and temporal correlations. The results did not account for the case of density-independence ($1-\gamma = 1$ or $\gamma = 0$).

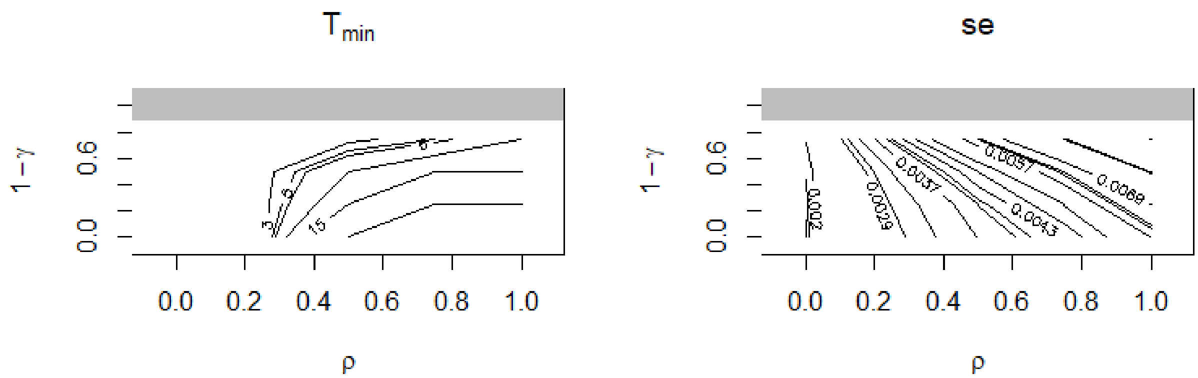

Fig 3 : Contour plots of the standard deviation of the trend estimate for the preferred number of temporal surveys and minimum sampling times as a function of temporal correlation $(1-\gamma)$ and spatial correlation (ρ) for only permanent plots. The survey cost is $B = 1500$, $\sigma^2 = 1$ and $\sigma^2_{obs} = 0.1$

Then we describe the second scenario which corresponded to the case where there were a mixture of permanent and non-permanent plots. The preferred strategy did also depend on spatial and temporal correlations. The results are the same for the optimal sampling time and the optimal standard deviation with a few differences. Indeed, when spatial correlation was high (ρ high) and temporal correlation low ($1-\gamma$ low), the optimal sampling times was $P_{min} = 15$ (the half of possible value). Therefore, allocating survey effort to temporal replicates was preferentially better (fig.4). Whereas when spatial correlation was low (ρ low) and temporal correlation high ($1-\gamma$ high), surveying in $T_{min} = 2$, the minimum of possible value, was the best strategy. That means that it would be better to allocate survey effort to spatial replicates (fig.4). Moreover, when spatial correlation was high (ρ high) and temporal correlation high ($1-\gamma$ high), surveying in $T_{min} = 2$ was also the best strategy (fig.4). The standard deviation of the estimated trend for the best strategy was lowest when spatial and temporal correlations were low and highest when spatial and temporal correlations were high (fig.4).

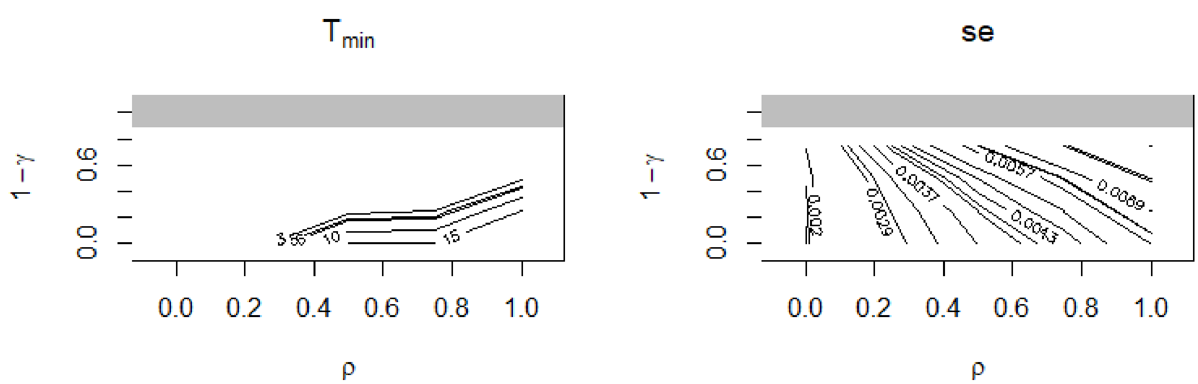

Fig 4 : Contour plots of the standard deviation of the trend estimate for the preferred number of temporal surveys and average minimum sampling times as a function of temporal correlation $(1-\gamma)$ and spatial correlation (ρ) for mixture of permanent and non-permanent plots. The survey cost is $B = 1500$, $\sigma^2 = 1$ and $\sigma^2_{obs} = 0.1$

We also optimize the proportion of non-permanents plots, p. So, it comes out that when spatial correlation was low (ρ low), the optimal sampling was $P_{min} = 0$. That means that the optimal sampling was the one without non-permanent plots (fig.5). Whereas when spatial correlation increased the proportion of non-permanent plots first increased and then decreased (fig.5) while knowing that the number of visits in the permanent plots increased regularly (fig.4). So, as the spatial correlation increased, the best strategy was to firstly sample with non-permanent plots and more temporal replicates in the temporal permanent plots. Then came a level where the best strategy was only permanent plots with temporal replicates. When temporal correlation was high (γ high), the change of strategy did not happen (fig.5). So, the best strategy in this case is sampling with non-permanent plots and more temporal replicates in the temporal permanent plots.

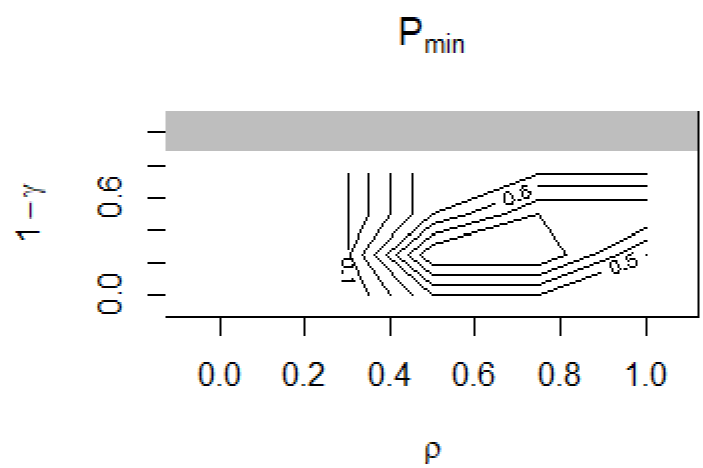

Fig 5 : Contour plots of proportion of non-permanents plots as a function of temporal correlation $(1-\gamma)$ and spatial correlation (ρ) for mixture of permanent and non-permanent plots. The survey cost is $B = 1500$, $\sigma^2 = 1$ and $\sigma^2_{obs} = 0.1$

After we compared the standard error of the estimator of trend, $\hat{r}$ in these two cases (presence and absence of non-permanent plots). The gain of standard error $(\Delta_{se} - 1)$ varied between 0 and 3%. When the spatial correlation was weak, the variation in the estimation error was equal to 1 (fig 6). Therefore there was no point in using sampling with non-permanent plots. While when spatial and temporal correlations are moderate, the best sampling was sampling with a proportion of non-permanent plots. The same conclusions were valid when temporal and spatial correlation were high. In this case, the gain of standard error (Δ_{se}-1) varied between 1 and 3% (fig 6).

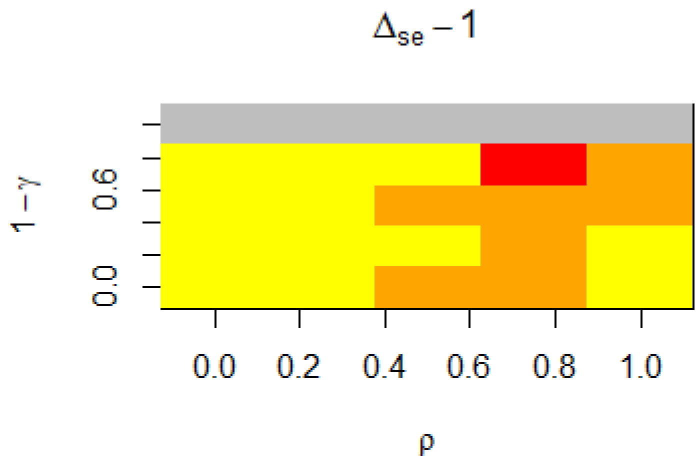

Fig 6 : Plots of the variation of standard deviation as a function of temporal correlation $(1 - \gamma)$ and spatial correlation (ρ). red is for $(\Delta_{se} - 1) \in] - 0.03, -0.02]$, orange is for $(\Delta_{se} - 1) \in] - 0.02, -0.01]$ and yellow is for $(\Delta_{se} - 1) \in] - 0.01, -0]$

5. Discussion

5.1. Is it necessary to introduce non-permanent plots in the sampling design?

Optimization of the average number of visits per site in presence/absence of non-permanent plots has shown us that when temporal correlation was low and spatial correlation was high, sampling with temporal replicates was the best strategy. Whereas when spatial correlation was low and temporal correlation is high, sampling with spatial replicates was the best strategy. These results are corroborated by the work of Rhodes and Jonsen (Rhodes and Jonzén 2011).

The spatial correlation reflects the existence of a similarity relationship between the information collected on two close sites at the same time t. While the temporal correlation reflects the similarity between the information collected at two different dates on the same site. Having a powerful sampling design need to have a maximum of independent observations. Correlation, whether temporal or spatial, violates this rule. In fact, when the correlation is higher, the amount of information contained in the sample is smaller (fig.5, right side). The estimator obtained in this case is therefore less efficient. In the presence of a strong spatial correlation, the information collected on non-permanent plots will be similar to that contained in permanent plots. The results obtained after optimizing the proportion of non-permanent plots seem to prove the contrary. Indeed, our analyzes show that when the spatial correlation was weak, introducing non-permanent plots did not provide additional information. This paradox is partly resolved when we remember the fact that the permanent plots actually had only two visits, so that there was already a lot of sites. The marginal gain in adding non-permanent plots was small and did not compensate for the loss in multiplying the between-site variability.

In order to measure the performance of the obtained temporal trend estimators, we calculated the standard error. The results obtained shown that the gain of estimating trend by introduction non-permanent plots was too weak (1-3%).

5.2. Limitations of the present results and associated perspectives

The first limit of our work is obviously the simplifying assumption of stationarity that we made for the simulation. Indeed, in real life, monitoring does not necessarily start in a equilibrium state. As Wikle and Royle affirmed in 1999 : *"assuming that one knows the spatio-temporal covariance structure of a process, it is simple to develop the best linear unbiased predictor and associated prediction variance for some location and time given a sample of observations. But generally, one does not know the full joint spatio-temporal covariance structure"*. Furthermore, the analysis of $\Delta_{se} - 1$ showed us that the gain in estimated trend is between 1 and 3% when we introduce non-permanent plots compared to sampling with only permanent plots. We therefore wonder if this gain is practically significant or not. So, we propose to complete the present work with implementing the complete non-stationary model through *Nimble* or *greta* and *TMB* (if it is possible). Secondly, the budget, the environmental variation and the variation of the observation error were constant. It would also be interesting to look for the effect of a variation of the budget in the optimal sampling method. The same for the cases where there is a variation of observation errors (σ^2_{obs}) and/or variation of σ^2_{trend}. In addition, due to the problem of access to the computing cluster, we were not able to perform the simulations in the case of the strata. As the monitoring project is national, even international and as terrestrial biodiversity in French is not homogeneous, it would be interesting to implement the model while introducing strata in order to take account of this heterogeneity. In ecology, the spatial and temporal correlations are not always independent as in our work. It would be interesting to model the dynamics while considering an interaction between temporal and spatial correlations, i.e. a truly spatio-temporal correlation structure. This will introduce a probably much greater complexity (see Wikle et al. 2019). In meta-population, they models are called *spatially-explicit metapopulation* models. Hanski et al. 1994 used this type of model to model the metapopulation structure and migration in the Butterfly Melitaea Cinxia.

6. Conclusion

This internship is part of the project called PASSIFOR2 (Proposals for the Improvement of the FOREST Biodiversity Monitoring System). It was funded by the Center for Ecology and Conservation Sciences (CESCO-Museum) and involved CESCO and INRAE Nogent-sur-Vernisson.
The objective of this internship is to optimize the distribution of the sampling effort in time and space based on Rhodes and Jonsen. Our work mainly consisted of clarify the limits and simplifying assumption of the original Rhodes and Jonzen model and propose necessary extensions for application in real managed ecosystems. After that we introduce the notion of non-permanent plots and strata in the sampling design. Then, we computed the non-stationary model on TMB and greta and the stationary one on R. After, we compared the results obtained in the last one with those of Rhodes and Jonsen. Due to time and the problem of accessibility to the computation cluster, all the simulations were not done. While implementing the non-stationary model, we encountered problems that have remained unresolved until now. The results obtained from the numerical simulation reflect the importance of the introduction of non-permanent plots in biodiversity monitoring. However, the quantitative gain is small (3% at most). This is potentially linked to our definition of permanent plots which takes into account plots with at least two passages. Many other research questions remain unanswered in this work. And it would be interesting to focus on it in order to improve our result in particular and biodiversity monitoring programs in general. We have listed a few above.

7. Appendix

7.1. Appendix A

We will first work on the probabilistic model. Then, we will develop the statistical model which will help to estimate the parameters. At this level, new parameters (initial conditions) will be introduced.

Inititial model

The Gompertz model can be written as :

$$N_{i,t} = N_{i,t-1}exp(-0.5\sigma^2 + \gamma(lnK_{i,t-1} - lnN_{i,t-1}) + u_{i,t}), \tag{A1}$$

Assumptions

- $lnK_{i,t}$ changes deterministically at a rate r (Rhodes and Jonsen's assumption 1), i.e $lnK_{i,t} = lnK_{i,t-1} + (r + \eta_i)$ (η_i appears if there is spatial variation of the trend)
- $Var(u_{i,t}) = \sigma^2$
- $Cov(u_{i,t}, u_{j,t}) = \rho^{d_{i,j}}\sigma^2$ (the growth rate of two different sub-populations, i and j, at the same point in time, t) with ρ the magnitude of spatial correlation in abundances, $d_{i,j}$ is the geographic distance between sub-populations i and j
- $0<\rho<1$ (the parameter that defines the strength of spatial correlation in growth rates)
- $0<\gamma<1$ (the parameter that defines the strength of temporal correlation and density-dependence. $\gamma = 0$ means density-independence whereas γ=1 means that there is not temporal correlation. In this case, the population abundance $lnN_{i,t}$ fluctuates around carrying capacity $lnK_{i,t-1}$)
- $Cov(u_{i,t}, u_{i,s}) = 0$ ($u_{i,t}$ are assumed to be independent with of $N_{i,0}$)
- $Cov(u_{i,t}, u_{j,s}) = 0$ for $i \neq j$ or $t \neq s$
- σ^2, γ and ρ are constant over time
- γ and ρ are assumed to be independent
- $Cov(u_{i,t}, \eta_i) = 0$ (environmental variation is assumed to not depend on variation in temporal trends of each sub-population)

Deriving the multivariate distribution of DFEs

On a logarithmic scale, the Gompertz model can be written as :

$$lnN_{i,t} = lnN_{i,t-1} - 0.5\sigma^2 - \gamma(lnK_{i,t-1} - lnN_{i,t-1}) + u_{i,t}$$

Rhodes and Jonsen define the quantity $\epsilon_{i,t} = lnN_{i,t} - lnK_{i,t}$ which we call here deviations from equilibrium (DFE). Then, $lnN_{i,t-1} = lnK_{i,t-1} + \epsilon_{i,t-1}$ and we can write the previous equation as

$$\begin{aligned} lnN_{i,t} &= lnK_{i,t-1} - 0.5\sigma^2 + \epsilon_{i,t-1} - \gamma\epsilon_{i,t-1} + u_{i,t} \\ &= lnK_{i,t-1} - 0.5\sigma^2 + (1-\gamma)\epsilon_{i,t-1} + u_{i,t} \end{aligned} \tag{A2}$$

As $lnK_{i,t}$ changes deterministically at a rate r (Rhodes and Jonsen assumption 1), i.e $lnK_{i,t} = lnK_{i,t-1} + r$, it holds that

$$lnK_{i,t-1} = lnK_{i,t} - r \tag{A3}$$

for $t \geq 1$.

By substituting eq.A3 into eq.A2, we obtain :

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - 0.5\sigma^2 \tag{A4}$$

In this work, we suppose that $\epsilon_{i,0}$ where $i \in \{1, ..., S\}$ with S the total number of sites, is multivariate gaussian variable. So, eq.A4 implies that $\epsilon_{i,t}$ with $i \in \{1, ..., S\}$ is a $S \times T$-dimensional gaussian vector, as a linear combination of $\epsilon_{i,0}$.
Now we can derive mean and variance-correlation structure of the deviations from local equilibrium, $\epsilon_{i,t}$.

Variance-covariance matrix of the models with deviation from equilibrium (DFE)

Mean of DFLE

$$\begin{aligned} Mean(\epsilon_{i,t}) &= Mean((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - 0.5\sigma^2) \\ &= (1-\gamma)Mean(\epsilon_{i,t-1}) - (r + 0.5\sigma^2) \end{aligned}$$

Using a general result about arithmetico-geometric sequences, we obtain :

$$Mean(\epsilon_{i,t}) = (1-\gamma)^t Mean(\epsilon_{i,0}) - (r + 0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{A5}$$

We can easily note that if $0<\gamma<1$, the sequence $Mean(\epsilon_{i,t})$ converges when t increases to infinity but is not constant. It converges when t becomes large ($t \to \infty$) and it is then equal to $\frac{-(r+0.5\sigma^2)}{\gamma}$. Having $Mean(\epsilon_{i,t})$ constant in time means that we must assume that it is initialized to its equilibrium value, i.e. $Mean(\epsilon_{i,0}) = \frac{-(r+0.5\sigma^2)}{\gamma}$. In this work, $0<\gamma<1$, that means that we always have convergence.

Variance

$$\begin{aligned} Var(\epsilon_{i,t}) &= Var((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - 0.5\sigma^2) \\ &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + \sigma^2 \end{aligned}$$

We find an arithmetico geometric sequence for which we have a standard method that gives us the general term of the sequence. Thus,

$$Var(\epsilon_{i,t}) = (1-\gamma)^{2t} Var(\epsilon_{i,0}) + \sigma^2 \frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2} \tag{A6}$$

If $0<\gamma<1$, asymptotically in t, we obtain $Var(\epsilon_{i,t}) \to \frac{\sigma^2}{1-(1-\gamma)^2}$ because $(1-\gamma)^{2t} \to 0$. That means that the sequence $Var(\epsilon_{i,t})$ converges when t increases to infinity but is not constant. Having $Var(\epsilon_{i,t})$ constant in time means that we must assume that it is initialized to its equilibrium value, i.e. $Var(\epsilon_{i,0}) = \frac{\sigma^2}{1-(1-\gamma)^2}$.

Covariance

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= Cov((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - 0.5\sigma^2, (1-\gamma)\epsilon_{j,t-s-1} + u_{j,t-s} - r - 0.5\sigma^2) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ &+ (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) \end{aligned} \tag{A7}$$

Based on the recurrence (eq.A4), we have $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)Cov(\epsilon_{i,t-2}, u_{j,t-s}) + Cov(u_{i,t-1}, u_{j,t-s})$.

We have $Cov(\epsilon_{i,t-v}, u_{j,t-s}) = 0$ if $v > s$. Both two previous results imply that

- if $s = 1$
  $Cov(\epsilon_{i,t-1}, u_{j,t-1}) = \rho^{d_{i,j}}\sigma^2$

- if $s > 1$
  $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)Cov(\epsilon_{i,t-2}, u_{j,t-s})$
  Then we can remark that $Cov(\epsilon_{i,t-s+v}, u_{j,t-s})$ is a geometric sequence on v where $(1-\gamma)$ is a common ratio and $\rho^{d_{i,j}}\sigma^2$ is the initial term. Therefore $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)^{s-1}\rho^{d_{i,j}}\sigma^2$.

Case 1 : $\mathbf{s > 0}$ $(t - s < t)$
We have $Cov(u_{i,t}, \epsilon_{j,t-s-1}) = 0$ and $Cov(u_{i,t}, u_{j,t-s}) = 0$. Then,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \end{aligned}$$

We find an arithmetico geometric sequence for which we have a standard method that gives us the general term of the sequence. Thus, we obtain

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2(t-s)} Cov(\epsilon_{i,s}, \epsilon_{j,0}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2}$$

In addition, $Cov(\epsilon_{i,s}, \epsilon_{j,0}) = (1-\gamma)Cov(\epsilon_{i,s-1}, \epsilon_{j,0}) + Cov(u_{i,s}, \epsilon_{j,0}) = (1-\gamma)Cov(\epsilon_{i,s-1}, \epsilon_{j,0})$ because $s > 0$. Using recurrence, we obtain

$$Cov(\epsilon_{i,s}, \epsilon_{j,0}) = (1-\gamma)^s Cov(\epsilon_{i,0}, \epsilon_{j,0}) \tag{A7a}$$

.

Then,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= (1-\gamma)^{2(t-s)}(1-\gamma)^s Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} \\ &= (1-\gamma)^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} \\ &= (1-\gamma)^{2t-s}(Cov(\epsilon_{i,0}, \epsilon_{j,0}) - \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}) + \frac{(1-\gamma)^s \rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \end{aligned}$$

Whence, asymptotically in t $(t \to \infty)$, with fixed s, we have $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^s \rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

Similarly, a simple change of variable $s = -s$, yields the result for $s < 0$ $(t - s > t)$ is $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s}(Cov(\epsilon_{i,0}, \epsilon_{j,0}) - \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}) + \frac{(1-\gamma)^{-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$. And asymptotically, we have : $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^{-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$

Case 2 : $\mathbf{s = 0}$
We have,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t}) \\ &\quad + (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-1}) + Cov(u_{i,t}, u_{j,t}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-1}) + Cov(u_{i,t}, u_{j,t}) \\ &= (1-\gamma)^{2t} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2} \\ &= (1-\gamma)^{2t} Cov(\epsilon_{i,0}, \epsilon_{j,0}) - \frac{(1-\gamma)^{2t}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} + \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \end{aligned}$$

because $Cov(\epsilon_{i,t-1}, u_{j,t}) = 0$ and $Cov(u_{i,t}, \epsilon_{j,t-1}) = 0$. Furthermore, $Cov(u_{i,t}, u_{j,t}) = \rho^{d_{i,j}}\sigma^2$. Then asymptotically in t with fixed s , we obtain $Cov(\epsilon_{i,t}, \epsilon_{j,t}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

On the whole, asymptotically in t with fixed s , we obtain

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \tag{A8}$$

As in the cases of mean and variance of $\epsilon_{i,t}$, if $0<\gamma<1$, the sequence $Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$ is not constant over time. Having $Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$ constant in time means that we must assumed following initial condition : $Cov(\epsilon_{i,0}, \epsilon_{j,0}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

If we sampled S subpopulations, each at T points in time, with no observation error, and β the number of years between two surveys, then, based on the equations above, the variance-covariance matrix of the deviations from equilibrium, $\epsilon_{i,t}$, is

$$\Phi = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} \tag{A9}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}} & (1-\gamma)^{\beta}\rho^{d_{i,j}} & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} \\ (1-\gamma)^{\beta}\rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}} \end{bmatrix}$$

Each sub-matrix $\mathbf{P}_{i,j}$ represents the correlation structure among the deviations from equilibrium for surveyed subpopulations i and j among points in time. To obtain above matrix, we assumed following initial conditions : $Mean(\epsilon_{i,0}) = \frac{-(r+0.5\sigma^2)}{\gamma}$, $Var(\epsilon_{i,0}) = \frac{\sigma^2}{1-(1-\gamma)^2}$ and $Cov(\epsilon_{i,0}, \epsilon_{j,0}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$ which amounts to assuming that the dynamics has been running for a long time before the beginning of the survey. If not, the variance-covariance matrix should be replaced by the non-asymptotic expression
$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s}(Cov(\epsilon_{i,0}, \epsilon_{j,0}) - \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}) + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

Variation in temporal trends among subpopulations

Here, let us suppose that each subpopulation has a different deterministic trend r_i, and that these trends are normally and independently distributed with mean r and variance σ^2_{trend}. Thus, the model is :

$$lnN_{i,t} = lnN_{i,t-1} - 0.5\sigma^2 - \gamma\epsilon_{i,t-1} + u_{i,t}$$

We call the quantity $\epsilon_{i,t} = lnN_{i,t} - lnK_{i,t} = lnN_{i,t} - lnK_{i,0} - (r+\eta_i)t$, deviations from local equilibrium (DFLE). η_i is an independent and normally distributed random variable with mean 0 and variance σ^2_{trend}. In this case, the DFLE is equal to

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - \eta_i - 0.5\sigma^2 \tag{A10}$$

Mean of DFLE
Based on the fact that $Mean(\eta_i) = 0$, we have

$$\begin{aligned} Mean(\epsilon_{i,t}) &= Mean((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - \eta_i - 0.5\sigma^2) \\ &= (1-\gamma)Mean(\epsilon_{i,t-1}) - (r+0.5\sigma^2) - Mean(\eta_i) \\ &= (1-\gamma)Mean(\epsilon_{i,t-1}) - (r+0.5\sigma^2) \end{aligned}$$

Then, using a general result about arithmetico-geometric sequences, we obtain

$$Mean(\epsilon_{i,t}) = (1-\gamma)^t Mean(\epsilon_{i,0}) - (r+0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{A11}$$

If $0<\gamma<1$, the sequence $Mean(\epsilon_{i,t})$ converges when t increases to infinity but is not constant. It converges when t becomes large ($t \to \infty$) and it is then equal to $\frac{-(r+0.5\sigma^2)}{\gamma}$. Having $Mean(\epsilon_{i,t})$ constant in time means that we must assume that it is initialized to its equilibrium value, i.e. $Mean(\epsilon_{i,0}) = \frac{-(r+0.5\sigma^2)}{\gamma}$.

Variance

$$\begin{aligned} Var(\epsilon_{i,t}) &= Var((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - \eta_i - 0.5\sigma^2) \\ &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + \sigma^2 + \sigma^2_{trend} - 2(1-\gamma)Cov(\epsilon_{i,t-1}, \eta_i) \end{aligned}$$

because $u_{i,t}$ and η_i are independent. We have

$$\begin{aligned} Cov(\epsilon_{i,t-1}, \eta_i) &= Cov((1-\gamma)\epsilon_{i,t-2} + u_{i,t} - r - \eta_i - 0.5\sigma^2), \eta_i) \\ &= (1-\gamma)Cov(\epsilon_{i,t-2}, \eta_i) - \sigma^2_{trend} \end{aligned}$$

We also find an arithmetico geometric sequence. By spreading the recurrence, we have :

$$Cov(\epsilon_{i,t-1}, \eta_i) = (1-\gamma)^{t-1} Cov(\epsilon_{i,0}, \eta_i) - \sigma^2_{trend}\frac{1-(1-\gamma)^{t-1}}{\gamma} \tag{A12}$$

Then,

$$\begin{aligned} Var(\epsilon_{i,t}) &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + \sigma^2 + \sigma^2_{trend} - 2((1-\gamma)^t Cov(\epsilon_{i,0}, \eta_i) - \sigma^2_{trend}\frac{(1-\gamma)-(1-\gamma)^t}{\gamma}) \\ &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + \sigma^2 + \sigma^2_{trend} + 2\sigma^2_{trend}\frac{(1-\gamma)}{\gamma} - 2(1-\gamma)^t (Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma}) \\ &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + B_1 - 2(1-\gamma)^t B_2 \end{aligned}$$

where $B_1 = \sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma})$ and $B_2 = Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma}$

By spreading the recurrence, we have :

$$Var(\epsilon_{i,t}) = (1-\gamma)^{2t} Var(\epsilon_{i,0}) + B_1\frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2} - 2B_2(1-\gamma)^t\frac{1-(1-\gamma)^t}{\gamma} \tag{A13}$$

Asymptotically in t, we obtain $Var(\epsilon_{i,t}) \to \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})) = \frac{B_1}{1-(1-\gamma)^2}$ because $(1-\gamma)^t \to 0$. If $0<\gamma<1$, the sequence $Var(\epsilon_{i,t})$ converges when t increases to infinity but is not constant. To obtain this constancy, we assumed following initial conditions : $Var(\epsilon_{i,0}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$ and $Cov(\epsilon_{i,0}, \eta_i) = -\frac{\sigma^2_{trend}}{\gamma}$ which amounts to assuming that the dynamics has been running for a long time before the beginning of the survey. These initial conditions are well defined because $Var(\epsilon_{i,0}) \geq -\frac{\sigma^2_{trend}}{\gamma}$ which can be demonstrated (proof not shown).

Covariance

Assuming that η_i and $u_{i,t}$ are independent, we have

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= Cov((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r - \eta_i - 0.5\sigma^2, (1-\gamma)\epsilon_{j,t-s-1} + u_{j,t-s} - r - \eta_j - 0.5\sigma^2) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) - (1-\gamma)Cov(\epsilon_{i,t-1}, \eta_j) \\ &+ (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) - (1-\gamma)Cov(\eta_i, \epsilon_{j,t-s-1}) + Cov(\eta_i, \eta_j) \end{aligned} \tag{A14}$$

Based on the recurrence (eq.A10), we have $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)Cov(\epsilon_{i,t-2}, u_{j,t-s}) + Cov(u_{i,t-1}, u_{j,t-s})$ because $Cov(\eta_i, u_{j,t-s}) = 0$.We have $Cov(\epsilon_{i,t-v}, u_{j,t-s}) = 0$ if $v > s$. Both two previous results imply that

- If $s = 1$
  $Cov(\epsilon_{i,t-1}, u_{j,t-1}) = \rho^{d_{i,j}}\sigma^2$
- If $s > 1$
  $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)Cov(\epsilon_{i,t-2}, u_{j,t-s})$
  Then we can remark that $Cov(\epsilon_{i,t-s+v}, u_{j,t-s})$ is a geometric sequence on v where $(1-\gamma)$ is a common ratio and $\rho^{d_{i,j}}\sigma^2$ is the initial term. Therefore $Cov(\epsilon_{i,t-1}, u_{j,t-s}) = (1-\gamma)^{s-1}\rho^{d_{i,j}}\sigma^2$.

Case 1 : $\mathbf{i \neq j}$
If $i \neq j$, $Cov(\eta_i, \eta_j) = 0$, $Cov(\epsilon_{i,t-1}, \eta_j) = 0$ and $Cov(\eta_i, \epsilon_{j,t-s-1}) = 0$. Then,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ + (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) \end{aligned}$$

- If $s > 0$ $(t - s < t)$
  We have $Cov(u_{i,t}, \epsilon_{j,t-s-1}) = 0$ and $Cov(u_{i,t}, u_{j,t-s}) = 0$. Then,

  $$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \end{aligned}$$

  Here we can find an arithmetico geometric sequence for which we have a standard method that gives us the general term of the sequence. In addition, as we proved above, $Cov(\epsilon_{i,s}, \epsilon_{j,0}) = (1-\gamma)^s Cov(\epsilon_{i,0}, \epsilon_{j,0})$. Then,

  $$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^s \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2}$$

  Asymptotically in t $(t \to \infty)$, with fixed s, we have $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^s \rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.
  Similarly, a simple change of variable $s = -s$, yields the result for $s < 0$ $(t - s > t)$ is $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^{-s}\rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2}$. And asymptotically, we have : $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^{-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$

- If $s = 0$
  We have,

  $$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t}) \\ &+ (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-1}) + Cov(u_{i,t}, u_{j,t}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-1}) + Cov(u_{i,t}, u_{j,t}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-1}) + \rho^{d_{i,j}}\sigma^2 \end{aligned}$$

  because $Cov(\epsilon_{i,t-1}, u_{j,t}) = 0$ and $Cov(u_{i,t}, \epsilon_{j,t-1}) = 0$. Then, we have

  $$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + \rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2}$$

  Asymptotically in t $(t \to \infty)$, with fixed s, we have $Cov(\epsilon_{i,t}, \epsilon_{j,t}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

On the whole, for $i \neq j$, the non asymptotic expression is :

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} \tag{14a}$$

And asymptotically in t with fixed s, we obtain $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$. If $0<\gamma<1$, the sequence $Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$ is not constant over time. It converges when t becomes large $(t \to \infty)$ towards

$\frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$. Having $Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$ constant in time means that we must assumed following initial condition $Cov(\epsilon_{i,0}, \epsilon_{j,0}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

Case 2 : $\mathbf{i = j}$
If $i = j$, we have:

$$\begin{aligned}Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{i,t-s}) - (1-\gamma)Cov(\epsilon_{i,t-1}, \eta_i) \\ &+ (1-\gamma)Cov(u_{i,t}, \epsilon_{i,t-s-1}) + Cov(u_{i,t}, u_{i,t-s}) - (1-\gamma)Cov(\eta_i, \epsilon_{i,t-s-1}) + \sigma^2_{trend}\end{aligned}$$

$Cov(\epsilon_{i,t-1}, \eta_i) = (1-\gamma)Cov(\epsilon_{i,t-2}, \eta_i) - \sigma^2_{trend}$ because $u_{i,t}$ and η_i are independent. According to eq.A12, we obtain :

$$Cov(\epsilon_{i,t-1}, \eta_i) = (1-\gamma)^{t-1}Cov(\epsilon_{i,0}, \eta_i) - \sigma^2_{trend}\frac{1-(1-\gamma)^{t-1}}{\gamma}$$

Similarly,

$$Cov(\eta_i, \epsilon_{i,t-s-1}) = (1-\gamma)^{t-s-1}Cov(\epsilon_{i,0}, \eta_i) - \sigma^2_{trend}\frac{1-(1-\gamma)^{t-s-1}}{\gamma}$$

- If $s > 0$ $(t - s < t)$
  We have $Cov(u_{i,t}, \epsilon_{i,t-s-1}) = 0$ and $Cov(u_{i,t}, u_{i,t-s}) = 0$. Then,

$$\begin{aligned}Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{i,t-s}) - (1-\gamma)Cov(\epsilon_{i,t-1}, \eta_i) \\ &- (1-\gamma)Cov(\eta_i, \epsilon_{i,t-s-1}) + \sigma^2_{trend} \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + (1-\gamma)^s\sigma^2 + \sigma^2_{trend} - (1-\gamma)^t Cov(\epsilon_{i,0}, \eta_i) + \sigma^2_{trend}\frac{(1-\gamma)-(1-\gamma)^t}{\gamma} \\ &- (1-\gamma)^{t-s}Cov(\epsilon_{i,0}, \eta_i) + \sigma^2_{trend}\frac{(1-\gamma)-(1-\gamma)^{t-s}}{\gamma} \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + (1-\gamma)^s\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) \\ &- (1-\gamma)^t(Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma}) - (1-\gamma)^{t-s}(Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + (1-\gamma)^s\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) \\ &- (1-\gamma)^t(Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1 + (1-\gamma)^{-s}) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{i,t-s-1}) + A_1 - A_2(1-\gamma)^t\end{aligned}$$

  where $A_1 = (1-\gamma)^s\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma})$ and $A_2 = (Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1 + (1-\gamma)^{-s})$
  By spreading the recurrence, we have :

$$Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) = (1-\gamma)^{2(t-s)}Cov(\epsilon_{i,s}, \epsilon_{j,0}) + A_1\frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} - A_2(1-\gamma)^t\frac{1-(1-\gamma)^{t-s}}{\gamma}$$

  As we showed previously in eq.A7a, $Cov(\epsilon_{i,s}, \epsilon_{j,0}) = (1-\gamma)^s Var(\epsilon_{i,0})$
  Thus,

$$\begin{aligned}Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) &= (1-\gamma)^{2t-s}Var(\epsilon_{i,0}) - (1-\gamma)^{2(t-s)}Cov(\epsilon_{i,0}, \eta_i)\frac{1-(1-\gamma)^s}{\gamma} + (1-\gamma)^s\sigma^2\frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} \\ &+ \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma})\frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} - (Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1 + (1-\gamma)^{-s})(1-\gamma)^t\frac{1-(1-\gamma)^{t-s}}{\gamma}\end{aligned}$$

Asymptotically in t ($t \to \infty$), with fixed s, we have

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{1}{1-(1-\gamma)^2}((1-\gamma)^s \sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$$

Similarly, a simple change of variable $s = -s$, yields the result for $s < 0$ ($t - s > t$) is $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \to \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{-s}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$

- If $s = 0$
  Based on the previous results,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) &= (1-\gamma)^{2t} Var(\epsilon_{i,0}) + (\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))\frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2} \\ &- 2(Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1-\gamma)^t \frac{1-(1-\gamma)^t}{\gamma} \end{aligned}$$

For $i = j$, the non-asymptotic expression is :

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) &= (1-\gamma)^{2t-s} Var(\epsilon_{i,0}) - (1-\gamma)^{2min(t,t-s)} Cov(\epsilon_{i,0}, \eta_i)\frac{1-(1-\gamma)^{|s|}}{\gamma} \\ &+ (1-\gamma)^s \sigma^2 \frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma})\frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} \\ &- (Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1 + (1-\gamma)^{-s})(1-\gamma)^t \frac{1-(1-\gamma)^{min(t,t-s)}}{\gamma} \end{aligned} \tag{14b}$$

Asymptotically in t ($t \to \infty$), with fixed s, $Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$. If $0<\gamma<1$, the sequence $Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$ converges when t increases to infinity but is not constant. One can obtain a constant sequence by setting the initial conditions : $Var(\epsilon_{i,0}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$ and $Cov(\epsilon_{i,0}, \eta_i) = \frac{-\sigma^2_{trend}}{\gamma}$.

On the whole, asymptotically in t, with fixed s, we have :

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) \times \delta_{i,j}) \tag{A15}$$

with $\delta_{i,j}$, Kronecker symbol.

The variance-covariance matrix of the DFLE, with β the number of years between two surveys, is

$$\begin{aligned} \mathbf{\Pi} &= \mathbf{\Phi} + \mathbf{\Phi}_{trend} \\ &= \frac{\sigma^2}{1-(1-\gamma)^2} \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) \begin{bmatrix} \mathbf{J} & 0 & \dots & 0 \\ 0 & \mathbf{J} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{J} \end{bmatrix} \end{aligned} \tag{A16}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}}\sigma^2 & (1-\gamma)^\beta \rho^{d_{i,j}}\sigma^2 & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 \\ (1-\gamma)^\beta \rho^{d_{i,j}}\sigma^2 & \rho^{d_{i,j}}\sigma^2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 & \dots & \dots & \rho^{d_{i,j}}\sigma^2 \end{bmatrix}$$

and

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 1 & \dots & \dots & 1 \end{bmatrix}$$

with $\mathbf{J}$, a matrix of dimension $T \times T$

To obtain above matrix, we assumed that the dynamics has been running for a long time before the beginning of the survey. If not, the variance-covariance matrix should be replaced by the non-asymptotic expression, i.e

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = \begin{cases} (1-\gamma)^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma)^{|s|} \rho^{d_{i,j}} \sigma^2 \frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} & \text{if } i \neq j \\ (1-\gamma)^{2min(t,t-s)} Var(\epsilon_{i,0}) - (1-\gamma)^{2min(t,t-s)} Cov(\epsilon_{i,0}, \eta_i) \frac{1-(1-\gamma)^{|s|}}{\gamma} + A_1 \frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} & \\ -A_2 (1-\gamma)^t \frac{1-(1-\gamma)^{min(t,t-s)}}{\gamma} & \text{if } i = j \end{cases} \tag{A17}$$

where $A_1 = (1-\gamma)^{|s|}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma})$ and $A_2 = (Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma^2_{trend}}{\gamma})(1 + (1-\gamma)^{-s})$

Building a statistical framework to estimate the trend on Ns

In this section, the goal is to model the log-linear temporal trend in the population sizes over time. To do so, we will use the models that can include a nonlinear structure in the mean, a precise parametrized form of variance-covariance, or even random effects whose parameters will have to be estimated. These different structures share the same parameters, which makes them difficult to estimate with classical statistical tools. So, the model must be written explicitly and estimated. We will consider both marginal and conditional forms of the model and estimate their parameters. Before expliciting toward the statistical models, we will clarify how we model the initial capacity parameter $lnK_{i,0}$ and the initial DFEs (or DFLEs), $\epsilon_{i,0}$. Here, we assume that the behavior of the initial population from start (stationary behavior) is the same as the asymptotic behavior of the population from probabilistic model. In this work, β is the number of years between two surveys.

How to model $lnK_{i,0}$ and $\epsilon_{i,0}$?

Modelisation of $lnK_{i,0}$

There are several ways to model $lnK_{i,0}$. Either we assume that they are deterministic or stochastic (with or without spatial correlation). In this work, we assume that $lnK_{i,0}$ are deterministic. More specifically, we suppose $lnK_{i,0}$ are correlated with another environmental variable called x_i that we know. Then, we have :

$$lnK_{i,0} = \mu + \omega x_i \tag{A18}$$

where μ and ω are unknown parameters that will have to be estimated.

Modelisation of $\epsilon_{i,0}$

As for the initial capacity parameter, it is necessary to model the multivariate initial DFEs (or DFLEs), $\epsilon_{i,0}$. It could also have several structures. But in this work, we assume that the structure of $\epsilon_{i,0}$ is non-stationary. Thus, we have :

- *Mean($\epsilon_{i,0}$)*

  We assume that Mean($\epsilon_{i,0}$) is a constant ($Mean(\epsilon_{i,0}) = c$), with $c \neq \frac{-(r+0.5\sigma^2)}{\gamma}$, an estimated parameter. It means that we consider that the populations are on average above or below the equilibrium at the beginning of the sampling, i.e they would initially be out of equilibrium, non-stationary, which seems important to us for a monitoring scheme.

- *Variance and covariance of $\epsilon_{i,0}$*
  Here, we assume that $Var(\epsilon_{i,0}) = \frac{\sigma_0^2}{1-(1-\gamma)^2}(\sigma^2 + \sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))$ and $Cov(\epsilon_{i,0},\epsilon_{j,0}) = \frac{\sigma_0^2}{1-(1-\gamma)^2}(\rho^{d_{i,j}}\sigma^2 + \sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))$ where $\sigma_0 > 0$, σ_{trend}, σ, γ and ρ are an estimated parameters. In addition, σ, γ and ρ are the same parameters as those used in environmental variation structure $u_{i,t}$.

- *Correlation between $\epsilon_{i,0}$ and η_i*
  In this work, we assume that there is a fixed correlation between $\epsilon_{i,0}$ and η_i. That means that $Corr(\epsilon_{i,0},\eta_i) = \frac{Cov(\epsilon_{i,0},\eta_i)\sqrt{1-(1-\gamma)^2}}{\sigma_0\sigma_{trend}\sqrt{(\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}} = \theta$. Then, we have $Cov(\epsilon_{i,0},\eta_i) = \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}$ with θ an unknown parameter that will have to be estimated.

If $\sigma_0 = 1$, $c = -\frac{r+0.5\sigma^2}{\gamma}$ and $\theta = -\frac{\sigma_{trend}\sqrt{1-(1-\gamma)^2}}{\gamma\sigma_0\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}$, we are in the stationary regime from start.

Now, we can recalculate mean and variance-covariance expressions in each of the previous cases. We will highlight the stationary and non-stationary structures. Let us remember that the **stationary state is the one where the mean and the variance of the DFLE (DFE) do not depend on t. That means that all the quantities take the asymptotic values, i.e $Mean(\epsilon_{i,t}) = -\frac{r+0.5\sigma^2}{\gamma}$, $Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))$, $Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})\times\delta_{i,j})$ and $Cov(\epsilon_{i,0},\eta_i) = -\frac{\sigma_{trend}^2}{\gamma}$.**

Reminder from the probabilistic model

Model without variations in temporal trends
In this case, $\sigma_{trend} = 0$. Based on to the previous results (eq.A5, A6, A7 and A8), we have :

- *Mean*

$$Mean(\epsilon_{i,t}) = c(1-\gamma)^t - (r+0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{A19}$$

- *Variance*

$$Var(\epsilon_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{2t}(\sigma_0^2-1)+1) \tag{A20}$$

- *Covariance*

$$Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{(1-\gamma)^{2t-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2-1) + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \tag{A21}$$

  Asymptotically in t, we have :
  $Mean(\epsilon_{i,t}) = \frac{-(r+0.5\sigma^2)}{\gamma}$, $Var(\epsilon_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}$ and $Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

In presence of observation error, we have :

$$Mean(\lambda_{i,t}) = c(1-\gamma)^t - (r+0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} - 0.5\sigma_{obs}^2 \tag{A22}$$

$$Var(\lambda_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{2t}(\sigma_0^2-1)+1) + \sigma_{obs}^2 \tag{A23}$$

$$Cov(\lambda_{i,t},\lambda_{j,t-s}) = Cov(\epsilon_{i,t},\epsilon_{j,t-s}) \tag{A24}$$

Asymptotically in t with fixed s, we have :
$Mean(\lambda_{i,t}) = \frac{-(r+0.5\sigma^2)}{\gamma} - 0.5\sigma_{obs}^2$, $Var(\lambda_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2} + \sigma_{obs}^2$ and $Cov(\lambda_{i,t},\lambda_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$.

Model with variations in temporal trends

Based on to the previous results (eq.A11, A13, A14a, A14b and A17), we have :

- *Mean*

$$Mean(\epsilon_{i,t}) = c(1-\gamma)^t - (r+0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{A25}$$

- *Variance*

$$\begin{aligned} Var(\epsilon_{i,t}) = {} & \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))((1-\gamma)^{2t}(\sigma_0^2-1)+1) \\ & - 2(\frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} + \frac{\sigma_{trend}^2}{\gamma})\frac{(1-\gamma)^t-(1-\gamma)^{2t}}{\gamma} \end{aligned} \tag{A26}$$

- *Covariance*
  When $i \neq j$, we have :

$$Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{(1-\gamma)^{2t-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2-1) + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \tag{A27}$$

  When $i = j$, we have :

$$\begin{aligned} Cov(\epsilon_{i,t},\epsilon_{i,t-s}) = {} & \frac{\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}{1-(1-\gamma)^2}((1-\gamma)^{2min(t,t-s)}(\sigma_0^2(1-\gamma)^{|s|}-1)+1) + \frac{(1-\gamma)^{2t-s}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2-1) + \frac{(1-\gamma)^{|s|}\sigma^2}{1-(1-\gamma)^2} \\ & + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} \times \frac{(1-\gamma)^{2t-s}-(1-\gamma)^{2min(t,t-s)}}{\gamma} \\ & + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} \times \frac{(1+(1-\gamma)^{-s})((1-\gamma)^{t\times min(t,t-s)}-(1-\gamma)^t)}{\gamma} \\ & - \sigma_{trend}^2\frac{(1+(1-\gamma)^{-s})((1-\gamma)^t-(1-\gamma)^{2min(t,t-s)})}{\gamma^2} \end{aligned} \tag{A28}$$

  Asymptotically in t with fixed s, we have :
  $Mean(\epsilon_{i,t}) = \frac{-(r+0.5\sigma^2)}{\gamma}$ and $Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))$ and $Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})\times\delta_{i,j})$.

In presence of observation error, we have :

$$Mean(\lambda_{i,t}) = c(1-\gamma)^t - (r+0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} - 0.5\sigma_{obs}^2 \tag{A29}$$

$$\begin{aligned} Var(\lambda_{i,t}) = {} & \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma_{obs}^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))((1-\gamma)^{2t}(\sigma_0^2-1)+1) \\ & - 2(\frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} + \frac{\sigma_{trend}^2}{\gamma})\frac{(1-\gamma)^t-(1-\gamma)^{2t}}{\gamma} \end{aligned} \tag{A30}$$

and

$$Cov(\lambda_{i,t}, \lambda_{j,t-s}) = Cov(\epsilon_{i,t}, \epsilon_{i,t-s}) \tag{A31}$$

Asymptotically in t with fixed s, we have :
$Mean(\lambda_{i,t}) = \frac{-(r+0.5\sigma^2)}{\gamma} - 0.5\sigma^2_{obs}$, $Var(\lambda_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{obs} + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))$ and $Cov(\lambda_{i,t}, \lambda_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) \times \delta_{i,j})$.

Regression models without temporal variation of trends

Here there is not difference between marginal and conditional forms because the model do not contains sites effects. We have :

$$\begin{aligned} lnN_{i,t} &= lnK_{i,0} + rt + \epsilon_{i,t} \\ &= \mu + \omega x_i + rt + \epsilon_{i,t} \end{aligned}$$

where $\epsilon_{i,t}$ is a normally distributed multivariate variable with variance matrix Φ and $Mean(\epsilon_{i,t}) = c(1-\gamma)^t - (r + 0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma}$. Thus, non linear regression model is :

$$\begin{aligned} lnN_{i,t} &= \mu + \omega x_i + Mean(\epsilon_{i,t}) + rt + \epsilon'_{i,t} \\ &= \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + (1-\gamma)^t(c + \frac{r + 0.5\sigma^2}{\gamma}) + \epsilon'_{i,t} \end{aligned} \tag{A32}$$

with $\epsilon'_{i,t} = \epsilon_{i,t} - Mean(\epsilon_{i,t})$ a centered multivariate gaussian variable. In this model, $-\frac{r+0.5\sigma^2}{\gamma}$ and r are respectively the intercept and slope of the linear term t, $(1-\gamma)^t(c + \frac{r+0.5\sigma^2}{\gamma})$ a non linear term in t and $\epsilon'_{i,t}$, the residuals. In the stationary regime, this model becomes linear because $c = -\frac{r+0.5\sigma^2}{\gamma}$:

$$lnN_{i,t} = \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + \epsilon'_{i,t}$$

r (a deterministic trend in the equilibrium population sizes), γ (the intensity of temporal dependence), ρ (the intensity of spatial correlation), σ (environmental variation), μ, ω, c, θ and σ_0 are the parameters that will have to be estimated in this model.
Consequently the variance-covariance matrix $\mathbf{\Phi}$ of the residuals is :

Stationary form
According to eq.A20 and A21, the stationary form of $\mathbf{\Phi}$ is :

$$\mathbf{\Phi} = \frac{\sigma^2}{1-(1-\gamma)^2} \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} \tag{A33}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}} & (1-\gamma)^\beta \rho^{d_{i,j}} & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} \\ (1-\gamma)^\beta \rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}} \end{bmatrix}$$

Non stationary form
Referring to eq.A20 and A21, the non stationary form of $\mathbf{\Phi}$ is :

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} \tag{A34}$$

with

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,1} & \dots & \zeta_{i,j}^{0,T-1} \\ \zeta_{i,j}^{1,0} & \zeta_{i,j}^{1,1} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{T-1,0} & \dots & \dots & \zeta_{i,j}^{T-1,T-1} \end{bmatrix}$$

where $\zeta_{i,j}^{t_1,t_2} = \frac{(1-\gamma)^{t_1+t_2}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2 - 1) + \frac{(1-\gamma)^{|t_1-t_2|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$

Model with observation error
Let us introduce independent and normally distributed observation errors with variance σ_{obs}^2 in order to account for observation error. We assume that observation errors are independent ($Cov(v_{i,t}, v_{j,s}) = 0$) and that ecological deviations from equilibrium and observation errors are independent ($Cov(\epsilon_{i,t}, v_{i,t}) = 0$). One may want to introduce that additional observation error term in the statistical model as follows :

Stationary form

$$\mathbf{\Phi} = \frac{\sigma^2}{1-(1-\gamma)^2} \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} + \sigma_{obs}^2 \begin{bmatrix} \mathbf{I}_T & 0 & \dots & 0 \\ 0 & \mathbf{I}_T & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{I}_T \end{bmatrix} \tag{A35}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}} & (1-\gamma)^{\beta}\rho^{d_{i,j}} & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} \\ (1-\gamma)^{\beta}\rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}} \end{bmatrix}$$

Non stationary form

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} + \sigma_{obs}^2 \begin{bmatrix} \mathbf{I}_T & 0 & \dots & 0 \\ 0 & \mathbf{I}_T & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{I}_T \end{bmatrix} \tag{A36}$$

with

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,1} & \dots & \zeta_{i,j}^{0,T-1} \\ \zeta_{i,j}^{1,0} & \zeta_{i,j}^{1,1} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{T-1,0} & \dots & \dots & \zeta_{i,j}^{T-1,T-1} \end{bmatrix}$$

where $\zeta_{i,j}^{t_1,t_2} = \frac{(1-\gamma)^{t_1+t_2}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2 - 1) + \frac{(1-\gamma)^{|t_1-t_2|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$

Modeling of $Y_{i,t} = lnN_{i,t} - lnN_{i,0}$
Instead of modeling $lnN_{i,t}$ over time, it is also possible to model $Y_{i,t} = lnN_{i,t} - lnN_{i,0}$. This model may

have the advantage to give a better estimation of the slope (r in our case). The number of parameters to be estimated is less than the previous model because here we do not model $lnK_{i,0}$. Thus, we have

$$\begin{aligned}
Y_{i,t} &= lnN_{i,t} - lnN_{i,0} \\
&= lnK_{i,0} + rt + \epsilon_{i,t} - lnK_{i,0} - \epsilon_{i,0} \\
&= Mean(\epsilon_{i,t} - \epsilon_{i,0}) + rt + \tau_{i,t} \\
&= -(\frac{r + 0.5\sigma^2}{\gamma} + c) + rt + (1-\gamma)^t(c + \frac{r + 0.5\sigma^2}{\gamma}) + \tau_{i,t}
\end{aligned} \tag{A37}$$

where $\tau_{i,t} = \epsilon_{i,t} - \epsilon_{i,0} - Mean(\epsilon_{i,t} - \epsilon_{i,0}) = \epsilon'_{i,t} - \epsilon'_{i,0}$ is a normally distributed multivariate variable with unknown variance-covariance matrix and $Mean(\tau_{i,t}) = 0$. In this model, the intercept is $-(\frac{r+0.5\sigma^2}{\gamma} + c)$ while r is the slope of the linear term t, $(1-\gamma)^t(c + \frac{r+0.5\sigma^2}{\gamma})$ a non linear term and $\tau_{i,t}$, the residuals. In the stationary regime ($c = -\frac{r+0.5\sigma^2}{\gamma}$), this model becomes linear without intercept :

$$Y_{i,t} = rt + \tau_{i,t} \tag{A37a}$$

Let derive variance and covariance of $\tau_{i,t}$.

Variance

$$\begin{aligned}
Var(\tau_{i,t}) &= Var(\epsilon_{i,t}) + Var(\epsilon_{i,0}) - 2Cov(\epsilon_{i,t}, \epsilon_{i,0}) \\
&= Var(\epsilon_{i,t}) + Var(\epsilon_{i,0}) - 2(1-\gamma)^t Var(\epsilon_{i,0}) \\
&= Var(\epsilon_{i,t}) + Var(\epsilon_{i,0})(1 - 2(1-\gamma)^t) \\
&= \frac{\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{2t}(\sigma_0^2 - 1) + 1) + \frac{\sigma^2\sigma_0^2}{1-(1-\gamma)^2}(1 - 2(1-\gamma)^t) \\
&= \frac{\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^{2t} + 1 - 2(1-\gamma)^t) - (1-\gamma)^{2t} + 1) \\
&= \frac{\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^t - 1)^2 - (1-\gamma)^{2t} + 1) \\
&= \frac{\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^t - 1)(\sigma_0^2((1-\gamma)^t - 1) - ((1-\gamma)^t + 1))
\end{aligned} \tag{A38}$$

In the stationary regime, we obtain $Var(\tau_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2 + 1)$.

Covariance

$$\begin{aligned}
Cov(\tau_{i,t}, \tau_{j,t-s}) &= Cov(\epsilon_{i,t} - \epsilon_{i,0}, \epsilon_{j,t-s} - \epsilon_{j,0}) \\
&= Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) - Cov(\epsilon_{i,t}, \epsilon_{j,0}) - Cov(\epsilon_{i,0}, \epsilon_{j,t-s}) + Cov(\epsilon_{i,0}, \epsilon_{j,0}) \\
&= Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) - (1-\gamma)^t Cov(\epsilon_{i,0}, \epsilon_{j,0}) - (1-\gamma)^{t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + Cov(\epsilon_{i,0}, \epsilon_{j,0}) \\
&= Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) + Cov(\epsilon_{i,0}, \epsilon_{j,0})(1 - (1-\gamma)^t - (1-\gamma)^{t-s})
\end{aligned}$$

As we showed previously,

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma)^{2t-s}(Cov(\epsilon_{i,0}, \epsilon_{j,0}) - \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}) + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$$

Then, we have :

$$\begin{aligned}
Cov(\tau_{i,t},\tau_{j,t-s}) &= (1-\gamma)^{2t-s}(Cov(\epsilon_{i,0},\epsilon_{j,0}) - \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}) + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} - Cov(\epsilon_{i,0},\epsilon_{j,0})((1-\gamma)^t + (1-\gamma)^{t-s} - 1) \\
&= Cov(\epsilon_{i,0},\epsilon_{j,0})((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1) - \frac{(1-\gamma)^{2t-s}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} + \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} \\
&= Cov(\epsilon_{i,0},\epsilon_{j,0})((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1) + \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{|s|} - (1-\gamma)^{2t-s}) \\
&= \frac{\rho^{d_{i,j}}\sigma^2\sigma_0^2}{1-(1-\gamma)^2}((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1) + \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{|s|} - (1-\gamma)^{2t-s}) \\
&= \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1) - (1-\gamma)^{2t-s} + (1-\gamma)^{|s|})
\end{aligned} \tag{A39}$$

- If $i = j$

$$Cov(\tau_{i,t},\tau_{j,t-s}) = \frac{\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1) - (1-\gamma)^{2t-s} + (1-\gamma)^{|s|})$$

- If $i \neq j$

$$Cov(\tau_{i,t},\tau_{j,t-s}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^{2t-s} - (1-\gamma)^t - (1-\gamma)^{t-s} + 1)) - (1-\gamma)^{2t-s} + (1-\gamma)^{|s|})$$

  In the stationary regime, we obtain $Cov(\tau_{i,t},\tau_{j,t-s}) = \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(1+(1-\gamma)^{|s|})+(1-(1-\gamma)^t-(1-\gamma)^{t-s})$.

Denoting $\mathbf{\Psi}$, the variance-covariance matrix of $\tau_{i,t}$ and according to eq.A35,A36 and A43, we can deduce that :

Stationary form
The stationary form of $\mathbf{\Psi}$ is :

$$\begin{aligned}
\mathbf{\Psi} &= \mathbf{\Phi} + \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S} \end{bmatrix} \\
&= \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S} \end{bmatrix}
\end{aligned} \tag{A40}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}} & (1-\gamma)^{\beta}\rho^{d_{i,j}} & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} \\ (1-\gamma)^{\beta}\rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}} \end{bmatrix}$$

and

$$\mathbf{Q}_{i,j} = \begin{bmatrix} \chi_{i,j}^{0,0} & \chi_{i,j}^{0,\beta} & \dots & \chi_{i,j}^{0,\beta(T-1)} \\ \chi_{i,j}^{\beta,0} & \chi_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \chi_{i,j}^{\beta(T-1),0} & \dots & \dots & \chi_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

with $\chi_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2})$

Non stationary form
The non stationary form of $\mathbf{\Psi}$ is :

$$\begin{aligned}
\mathbf{\Psi} &= \mathbf{\Phi} + \frac{\sigma^2\sigma_0^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma^2\sigma_0^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S} \\ \dots & \dots & \dots & \dots \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S} \end{bmatrix}
\end{aligned} \tag{A41}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,\beta} & \dots & \zeta_{i,j}^{0,\beta(T-1)} \\ \zeta_{i,j}^{\beta,0} & \zeta_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{\beta(T-1),0} & \dots & \dots & \zeta_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

and

$$\mathbf{Q}_{i,j} = \begin{bmatrix} \chi_{i,j}^{0,0} & \chi_{i,j}^{0,\beta} & \dots & \chi_{i,j}^{0,\beta(T-1)} \\ \chi_{i,j}^{\beta,0} & \chi_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \chi_{i,j}^{\beta(T-1),0} & \dots & \dots & \chi_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

with $\zeta_{i,j}^{t_1,t_2} = \frac{(1-\gamma)^{t_1+t_2}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2-1) + \frac{(1-\gamma)^{|t_1-t_2|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$ and $\chi_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2})$

Model with observation error
If there is an observation error, the model (eq.A37) becomes :

$$\begin{aligned}
Y_{i,t} &= ln\tilde{N}_{i,t} - ln\tilde{N}_{i,0} \\
&= lnN_{i,t} - 0.5\sigma_{obs}^2 + v_{i,t} - lnN_{i,0} + 0.5\sigma_{obs}^2 - v_{i,0} \\
&= lnN_{i,t} - lnN_{i,0} + v_{i,t} - v_{i,0} \\
&= rt + \epsilon_{i,t} - \epsilon_{i,0} + v_{i,t} - v_{i,0} \\
&= -(\frac{r+0.5\sigma^2}{\gamma} + c) + rt + (1-\gamma)^t(c + \frac{r+0.5\sigma^2}{\gamma}) + \tau'_{i,t}
\end{aligned} \tag{A42}$$

where $\tau'_{i,t} = \tau_{i,t} + v_{i,t} - v_{i,0}$ is a normally distributed multivariate variable with unknown variance-covariance matrix and $Mean(\tau'_{i,t}) = 0$.
Let derive variance and covariance of $\tau'_{i,t}$.

Variance

$$Var(\tau'_{i,t}) = Var(\tau_{i,t}) + Var(v_{i,t}) + Var(v_{i,0}) - 2Cov(v_{i,t}, v_{i,0})$$

Since,

$$Cov(v_{i,t}, v_{i,0}) = \begin{cases} 0 & \text{if } t > 0 \\ \sigma_{obs}^2 & \text{if } i = j \end{cases}$$

We have,

$$\begin{aligned} Var(\tau'_{i,t}) &= Var(\tau_{i,t}) + 2\sigma^2_{obs}\mathbb{I}_{t>0} \\ &= \frac{\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2((1-\gamma)^{2t}+1-2(1-\gamma)^t) - (1-\gamma)^{2t}+1) + 2\sigma^2_{obs}\mathbb{I}_{t>0} \end{aligned} \tag{A43}$$

because $Var(v_{i,t}) = Var(v_{i,0}) = \sigma^2_{obs}$

In the stationary regime, we obtain $Var(\tau'_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}(2-2(1-\gamma)^t) + 2\sigma^2_{obs}\mathbb{I}_{t>0}$.

Covariance

$$\begin{aligned} Cov(\tau'_{i,t}, \tau'_{j,t-s}) &= Cov(\tau_{i,t} + v_{i,t} - v_{i,0}, \tau_{j,t-s} + v_{j,t-s} - v_{j,0}) \\ &= Cov(\tau_{i,t}, \tau_{j,t-s}) + Cov(\tau_{i,t}, v_{j,t-s}) - Cov(\tau_{i,t}, v_{j,0}) + Cov(v_{i,t}, \tau_{j,t-s}) + Cov(v_{i,t}, v_{j,t-s}) \\ &\quad - Cov(v_{i,t}, v_{j,0}) - Cov(v_{i,0}, \tau_{j,t-s}) - Cov(v_{i,0}, v_{j,t-s}) + Cov(v_{i,0}, v_{j,0}) \end{aligned} \tag{A44}$$

Case 1 : $i \neq j$
If $i \neq j$, we have :

$$Cov(\tau'_{i,t}, \tau'_{j,t-s}) = Cov(\tau_{i,t}, \tau_{i,t-s})$$

because $Cov(v_{i,t}, v_{j,s}) = 0$ and $Cov(\tau_{i,t}, v_{j,t}) = 0$
In the stationary regime, we obtain $Cov(\tau'_{i,t}, \tau'_{j,t-s}) \rightarrow \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(1+(1-\gamma)^{|s|})$

Case 2 : $i = j$
If $i = j$, we have :

$$\begin{aligned} Cov(\tau'_{i,t}, \tau'_{i,t-s}) &= Cov(\tau_{i,t}, \tau_{i,t-s}) + Cov(\tau_{i,t}, v_{i,t-s}) - Cov(\tau_{i,t}, v_{i,0}) + Cov(v_{i,t}, \tau_{i,t-s}) + Cov(v_{i,t}, v_{i,t-s}) \\ &\quad - Cov(v_{i,t}, v_{i,0}) - Cov(v_{i,0}, \tau_{i,t-s}) - Cov(v_{i,0}, v_{i,t-s}) + \sigma^2_{obs} \end{aligned}$$

- If $s \neq 0$
  We have :

  $$Cov(\tau'_{i,t}, \tau'_{i,t-s}) = Cov(\tau_{i,t}, \tau_{i,t-s}) - Cov(v_{i,t}, v_{i,0}) - Cov(v_{i,0}, v_{i,t-s}) + \sigma^2_{obs}$$

  because $Cov(\tau_{i,t}, v_{i,t-s}) = 0$ and $Cov(v_{i,t}, v_{i,t-s}) = 0$ if $t \neq 0$.
  As we showed above, $Cov(v_{i,t}, v_{i,0}) = \sigma^2_{obs}\mathbb{I}_{t=0}$ and $Cov(v_{i,0}, v_{i,t-s}) = \sigma^2_{obs}\mathbb{I}_{t=s}$
  So,

  $$Cov(\tau'_{i,t}, \tau'_{i,t-s}) = Cov(\tau_{i,t}, \tau_{i,t-s}) + \sigma^2_{obs}(1 - \mathbb{I}_{t=0} - \mathbb{I}_{t=s})$$

  In the stationary regime, $Cov(\tau'_{i,t}, \tau'_{j,t-s}) \rightarrow \frac{\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}((1-\gamma)^{|s|} + (1-(1-\gamma)^t - (1-\gamma)^{t-s}) + \sigma^2_{obs}(1 - \mathbb{I}_{t=0} - \mathbb{I}_{t=s})$

- If $s = 0$
  We have :

  $$Cov(\tau'_{i,t}, \tau'_{i,t}) = Var(\tau'_{i,t})$$

  In the stationary regime, we have $Cov(\tau'_{i,t}, \tau'_{i,t}) = \frac{\sigma^2}{1-(1-\gamma)^2}(2-2(1-\gamma)^t) + 2\sigma^2_{obs}\mathbb{I}_{t>0}$.

Denoting $\mathbf{\Delta}$, the structure of the variance-covariance matrix of $\tau'_{i,t}$, we obtain :

Stationary form
The stationary form of $\boldsymbol{\Delta}$ is

$$\begin{aligned}\boldsymbol{\Delta} &= \boldsymbol{\Psi} + \sigma^2_{obs}\begin{bmatrix}\mathbf{K} & 0 & \dots & 0\\ 0 & \mathbf{K} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{K}\end{bmatrix}\\ &= \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S}\\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S}\end{bmatrix} + \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S}\\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S}\end{bmatrix} + \sigma^2_{obs}\begin{bmatrix}\mathbf{K} & 0 & \dots & 0\\ 0 & \mathbf{K} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{K}\end{bmatrix}\end{aligned} \tag{A45}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix}\rho^{d_{i,j}} & (1-\gamma)^{\beta}\rho^{d_{i,j}} & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\\ (1-\gamma)^{\beta}\rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}}\end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix}\rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \rho^{d_{i,j}}\\ \rho^{d_{i,j}} & \rho^{d_{i,j}} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ \rho^{d_{i,j}} & \dots & \dots & \rho^{d_{i,j}}\end{bmatrix}$$

and

$$\mathbf{K} = \begin{bmatrix}0 & 0 & \dots & 0\\ 0 & 2 & 1 & 1\\ \dots & 1 & 2 & \dots\\ 0 & 1 & \dots & 2\end{bmatrix}$$

with $\mathbf{K}$, a matrix f dimension $T \times T$

Non asymptotic form
The non stationary form of $\boldsymbol{\Delta}$ is

$$\begin{aligned}\boldsymbol{\Delta} &= \boldsymbol{\Psi} + \sigma^2_{obs}\begin{bmatrix}\mathbf{K} & 0 & \dots & 0\\ 0 & \mathbf{K} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{K}\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S}\\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S}\end{bmatrix} + \frac{\sigma^2\sigma_0^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S}\\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S}\end{bmatrix} + \sigma^2_{obs}\begin{bmatrix}\mathbf{K} & 0 & \dots & 0\\ 0 & \mathbf{K} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{K}\end{bmatrix}\end{aligned} \tag{A46}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix}\zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,\beta} & \dots & \zeta_{i,j}^{0,\beta(T-1)}\\ \zeta_{i,j}^{\beta,0} & \zeta_{i,j}^{\beta,\beta} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ \zeta_{i,j}^{\beta(T-1),0} & \dots & \dots & \zeta_{i,j}^{\beta(T-1),\beta(T-1)}\end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix}\chi_{i,j}^{0,0} & \chi_{i,j}^{0,\beta} & \dots & \chi_{i,j}^{0,\beta(T-1)}\\ \chi_{i,j}^{\beta,0} & \chi_{i,j}^{\beta,\beta} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ \chi_{i,j}^{\beta(T-1),0} & \dots & \dots & \chi_{i,j}^{\beta(T-1),\beta(T-1)}\end{bmatrix}$$

and

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 2 & 1 & 1 \\ \dots & 1 & 2 & \dots \\ 0 & 1 & \dots & 2 \end{bmatrix}$$

with
$\zeta_{i,j}^{t_1,t_2} = \frac{(1-\gamma)^{t_1+t_2}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}(\sigma_0^2 - 1) + \frac{(1-\gamma)^{|t_1-t_2|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$ and $\chi_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(1-(1-\gamma)^{t_1} - (1-\gamma)^{t_2})$

Regression models with temporal variation of trends

Marginal approach
In the case where there is variation in temporal trends among subpopulations, the regression model (eq.A32) becomes :

$$lnN_{i,t} = \mu + \omega x_i - \frac{r+0.5\sigma^2}{\gamma} + rt + (1-\gamma)^t(c + \frac{r+0.5\sigma^2}{\gamma}) + \epsilon'_{i,t} + \eta_i t \tag{A47}$$

Here the residuals become $E_{i,t} = \epsilon'_{i,t} + \eta_i t$. In the stationary regime, this model becomes linear because $c = -\frac{r+0.5\sigma^2}{\gamma}$ and we obtain

$$lnN_{i,t} = \mu + \omega x_i - \frac{r+0.5\sigma^2}{\gamma} + rt + E_{i,t}$$

r (a deterministic trend in the equilibrium population sizes), γ (the intensity of temporal dependence), ρ (the intensity of spatial correlation), σ (environmental variation), μ, ω, c, σ_{trend}, θ and σ_0 are the parameters that will have to be estimated in the non-stationary form of the model.

Now we can derive mean and variance-correlation structure of the residuals, $E_{i,t}$, for this model.

Mean of residuals
As $Mean(\epsilon'_{i,t}) = 0$ and $Mean(\eta_i) = 0$ then $Mean(E_{i,t}) = 0$

Variance of residuals

$$\begin{aligned} Var(E_{i,t}) &= Var(\eta_i t + \epsilon'_{i,t}) \\ &= t^2\sigma_{trend}^2 + Var(\epsilon'_{i,t}) + 2tCov(\eta_i, \epsilon_{i,t}) \\ &= t^2\sigma_{trend}^2 + Var(\epsilon_{i,t}) + 2tCov(\eta_i, \epsilon_{i,t}) \end{aligned}$$

Using the same recurrence in eq.A12, we have

$$Cov(\epsilon_{i,t}, \eta_i) = (1-\gamma)^t Cov(\epsilon_{i,0}, \eta_i) - \sigma_{trend}^2 \frac{1-(1-\gamma)^t}{\gamma}$$

Thus,

$$\begin{aligned} Var(E_{i,t}) &= Var(\epsilon'_{i,t}) + t^2\sigma_{trend}^2 - \frac{2t\sigma_{trend}^2}{\gamma} + 2t(1-\gamma)^t(Cov(\epsilon_{i,0}, \eta_i) + \frac{\sigma_{trend}^2}{\gamma}) \\ &= Var(\epsilon_{i,t}) + t^2\sigma_{trend}^2 - \frac{2t\sigma_{trend}^2}{\gamma} + 2t(1-\gamma)^t(\frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2 + \sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}} + \frac{\sigma_{trend}^2}{\gamma}) \end{aligned} \tag{A48}$$

because $Cov(\epsilon_{i,0}, \eta_i) = \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}$ and $Var(\epsilon'_{i,t}) = Var(\epsilon_{i,t})$
In the stationary regime, we have : $Var(\epsilon_{i,0}) = Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma}))$ and

$Cov(\epsilon_{i,0},\eta_i) = -\frac{\sigma^2_{trend}}{\gamma}$. Then, we obtain : $Var(E_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})) + t^2\sigma^2_{trend} - \frac{2t\sigma^2_{trend}}{\gamma}$

Covariance of residuals
We have

$$\begin{aligned} Cov(E_{i,t},E_{j,t-s}) &= Cov(\eta_i t + \epsilon_{i,t} - Mean(\epsilon_{i,t}), \eta_j(t-s) + \epsilon_{j,t-s} - Mean(\epsilon_{j,t-s})) \\ &= t(t-s)Cov(\eta_i,\eta_j) + tCov(\eta_i,\epsilon_{j,t-s}) + (t-s)Cov(\eta_j,\epsilon_{i,t}) + Cov(\epsilon_{i,t},\epsilon_{j,t-s}) \end{aligned} \tag{A49}$$

Case 1 : i = j
If $i=j$, we have :

$$\begin{aligned} Cov(E_{i,t},E_{j,t-s}) = Cov(\epsilon_{i,t},\epsilon_{i,t-s}) + \sigma^2_{trend}(t(t-s) - t\frac{1-(1-\gamma)^{t-s}}{\gamma} - (t-s)\frac{1-(1-\gamma)^t}{\gamma}) \\ + Cov(\epsilon_{i,0},\eta_i)(t(1-\gamma)^{t-s} + (t-s)(1-\gamma)^t) \end{aligned}$$

In the stationary regime, we have : $Cov(\epsilon_{i,t},\epsilon_{i,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\sigma^2 + \sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))$ and $Cov(\epsilon_{i,0},\eta_i) = -\frac{\sigma^2_{trend}}{\gamma}$. Then, we obtain : $Cov(E_{i,t},E_{i,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\sigma^2 + \sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})) + \sigma^2_{trend}(t(t-s) - \frac{2t-s}{\gamma})$

Case 2 : i ≠ j
If $i \neq j$, we have $Cov(\eta_i,\eta_j)=0$, $Cov(\eta_i,\epsilon_{j,t-s})=0$ and $Cov(\eta_j,\epsilon_{i,t})=0$.
So, $Cov(E_{i,t},E_{j,t-s}) = Cov(\epsilon_{i,t},\epsilon_{j,t-s})$. In the stationary regime, we have : $Cov(E_{i,t},E_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$

Denoting $\mathbf{\Lambda}$ the variance-covariance matrix of the residuals $E_{i,t}$, we can easily see that $\mathbf{\Lambda}$ is a sum of matrices :

Stationary form
By referring to eq.A16, we have :

$$\begin{aligned} \mathbf{\Lambda} &= \mathbf{\Lambda} + \sigma^2_{trend}\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix} \\ &= \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{J} & 0 & ... & 0 \\ 0 & \mathbf{J} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{J} \end{bmatrix} + \sigma^2_{trend}\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix} \end{aligned} \tag{A50}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}}\sigma^2 & (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 \\ (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \rho^{d_{i,j}}\sigma^2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 & \dots & \dots & \rho^{d_{i,j}}\sigma^2 \end{bmatrix}$$

and

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 1 & \dots & \dots & 1 \end{bmatrix}$$

and

$$\mathbf{C}_{i,j} = \begin{bmatrix} c_{i,j}^{0,0} & c_{i,j}^{0,\beta} & \dots & c_{i,j}^{0,\beta(T-1)} \\ c_{i,j}^{\beta,0} & c_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ c_{i,j}^{\beta(T-1),0} & \dots & \dots & c_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

with $c_{i,j}^{t_1,t_2} = t_1t_2 - \frac{t_1+t_2}{\gamma}$

Non asymptotic form
By referring to eq.A17, we obtain :

$$\mathbf{\Lambda} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma_0^2\sigma_{trend}^2}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & ... & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & ... & \mathbf{Q}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & ... & \mathbf{Q}_{S,S} \end{bmatrix} + \sigma_{trend}^2\begin{bmatrix} \mathbf{L}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{L}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{L}_{S,S} \end{bmatrix}$$
$$+ \frac{\sigma_{trend}^2}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix} + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}\begin{bmatrix} \mathbf{G}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{G}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{G}_{S,S} \end{bmatrix} \tag{A51}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,\beta} & \dots & \zeta_{i,j}^{0,\beta(T-1)} \\ \zeta_{i,j}^{\beta,0} & \zeta_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{\beta(T-1),0} & \dots & \dots & \zeta_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix} q_{i,j}^{0,0} & q_{i,j}^{0,\beta} & \dots & q_{i,j}^{0,\beta(T-1)} \\ q_{i,j}^{\beta,0} & q_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ q_{i,j}^{\beta(T-1),0} & \dots & \dots & q_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{L}_{i,j} = \begin{bmatrix} l_{i,j}^{0,0} & l_{i,j}^{0,\beta} & \dots & l_{i,j}^{0,\beta(T-1)} \\ l_{i,j}^{\beta,0} & l_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ l_{i,j}^{\beta(T-1),0} & \dots & \dots & l_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{C}_{i,j} = \begin{bmatrix} c_{i,j}^{0,0} & c_{i,j}^{0,\beta} & \dots & c_{i,j}^{0,\beta(T-1)} \\ c_{i,j}^{\beta,0} & c_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ c_{i,j}^{\beta(T-1),0} & \dots & \dots & c_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

and

$$\mathbf{G}_{i,j} = \begin{bmatrix} g_{i,j}^{0,0} & g_{i,j}^{0,\beta} & \dots & g_{i,j}^{0,\beta(T-1)} \\ g_{i,j}^{\beta,0} & g_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ g_{i,j}^{\beta(T-1),0} & \dots & \dots & g_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

with $\zeta_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}((1-\gamma)^{t_1+t_2}\sigma_0^2 + (1-\gamma)^{|t_1-t_2|}(1-(1-\gamma)^{t_1+t_2-|t_1-t_2|}))$, $q_{i,j}^{t_1,t_2} = (1-\gamma)^{t_1+t_2}$, $c_{i,j}^{t_1,t_2} = 1-(1-\gamma)^{t_1+t_2-|t_1-t_2|}$, $l_{i,j}^{t_1,t_2} = t_1t_2 - t_1\frac{1-(1-\gamma)^{t_2}}{\gamma} - t_2\frac{1-(1-\gamma)^{t_1}}{\gamma} - \frac{(1-\gamma)^{t_1}-(1-\gamma)^{t_2}}{\gamma} \times \frac{1-(1-\gamma)^{t_2}}{\gamma}$ and $g_{i,j}^{t_1,t_2} = t_1(1-\gamma)^{t_2} + t_2(1-\gamma)^{t_1} - \frac{(1-\gamma)^{t_1}-(1-\gamma)^{t_2}}{\gamma} \times \frac{1-(1-\gamma)^{t_2}}{\gamma} - \frac{(1-\gamma)^{t_1+t_2-|t_1-t_2|}-(1-\gamma)^{t_1+t_2}}{\gamma}$

Model with observation error
Here we assume that observation errors are not depend on variation in temporal trends of each sub-population ($Cov(v_{i,t},\eta_i) = 0$). One may want to introduce an additional observation error term in the statistical model as follows :

Stationary form

By referring to eq.A30, A31 and A50, we have :

$$\begin{aligned}\mathbf{\Lambda} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S}\\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S}\end{bmatrix} + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix}\mathbf{J} & 0 & \dots & 0\\ 0 & \mathbf{J} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{J}\end{bmatrix}\\ + \sigma^2_{trend}\begin{bmatrix}\mathbf{C}_{1,1} & 0 & \dots & 0\\ 0 & \mathbf{C}_{2,2} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{C}_{S,S}\end{bmatrix} + \frac{\sigma^2_{obs}}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{I}_T & 0 & \dots & 0\\ 0 & \mathbf{I}_T & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{I}_T\end{bmatrix}\end{aligned} \tag{A52}$$

Non stationary form

$$\begin{aligned}\mathbf{\Lambda} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S}\\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S}\\ ... & ... & ... & ...\\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S}\end{bmatrix} + \frac{\sigma_0^2\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix}\mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & ... & \mathbf{Q}_{1,S}\\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & ... & \mathbf{Q}_{2,S}\\ ... & ... & ... & ...\\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & ... & \mathbf{Q}_{S,S}\end{bmatrix} + \sigma^2_{trend}\begin{bmatrix}\mathbf{L}_{1,1} & 0 & ... & 0\\ 0 & \mathbf{L}_{2,2} & ... & 0\\ ... & ... & ... & ...\\ 0 & 0 & ... & \mathbf{L}_{S,S}\end{bmatrix}\\ + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix}\mathbf{C}_{1,1} & 0 & ... & 0\\ 0 & \mathbf{C}_{2,2} & ... & 0\\ ... & ... & ... & ...\\ 0 & 0 & ... & \mathbf{C}_{S,S}\end{bmatrix} + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}\begin{bmatrix}\mathbf{G}_{1,1} & 0 & ... & 0\\ 0 & \mathbf{G}_{2,2} & ... & 0\\ ... & ... & ... & ...\\ 0 & 0 & ... & \mathbf{G}_{S,S}\end{bmatrix}\\ + \frac{\sigma^2_{obs}}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{I}_T & 0 & ... & 0\\ 0 & \mathbf{I}_T & ... & 0\\ ... & ... & ... & ...\\ 0 & 0 & ... & \mathbf{I}_T\end{bmatrix}\end{aligned} \tag{A53}$$

Modeling of $Y_{i,t} = lnN_{i,t} - lnN_{i,0}$

As for the model without sites effects, we will also model $Y_{i,t} = lnN_{i,t} - lnN_{i,0}$. Based to eq.A37, we have

$$\begin{aligned}Y_{i,t} &= lnN_{i,t} - lnN_{i,0}\\ &= lnK_{i,0} + rt + \eta_i t + \epsilon_{i,t} - lnK_{i,0} - \epsilon_{i,0}\\ &= -(\frac{r+0.5\sigma^2}{\gamma} + c) + rt + (1-\gamma)^t(c + \frac{r+0.5\sigma^2}{\gamma}) + E'_{i,t}\end{aligned} \tag{A54}$$

where $E'_{i,t} = \tau_{i,t} + \eta_i t$ is a normally distributed multivariate variable with parametrized variance-covariance matrix and $Mean(E'_{i,t}) = 0$. In this model, the intercept is $-(\frac{r+0.5\sigma^2}{\gamma} + c)$ while r is the slope of the linear term t, $(1-\gamma)^t(c+\frac{r+0.5\sigma^2}{\gamma})$ a non linear term and $E'_{i,t}$, the residuals. In the stationary regime ($c = -\frac{r+0.5\sigma^2}{\gamma}$), this model becomes linear without intercept :

$$Y_{i,t} = rt + E'_{i,t} \tag{A54a}$$

Let us derive variance and covariance of $E'_{i,t}$.

Variance

$$\begin{aligned}Var(E'_{i,t}) &= Var(\tau_{i,t} + \eta_i t)\\ &= Var(\tau_{i,t}) + t^2\sigma^2_{trend} + 2tCov(\eta_i, \tau_{i,t})\end{aligned}$$

$$\begin{aligned}
Cov(\eta_i,\tau_{i,t}) &= Cov(\epsilon'_{i,t}-\epsilon'_{i,0},\eta_i)\\
&= Cov(\eta_i,\epsilon'_{i,t}) - Cov(\eta_i,\epsilon'_{i,0})\\
&= Cov(\eta_i,\epsilon_{i,t}) - Cov(\eta_i,\epsilon_{i,0})\\
&= (1-\gamma)^t Cov(\epsilon_{i,0},\eta_i) - \sigma^2_{trend}\frac{1-(1-\gamma)^t}{\gamma} - Cov(\epsilon_{i,0},\eta_i)\\
&= (1-\gamma)^t(Cov(\epsilon_{i,0},\eta_i)+\frac{\sigma^2_{trend}}{\gamma}) - (Cov(\epsilon_{i,0},\eta_i)+\frac{\sigma^2_{trend}}{\gamma})\\
&= (Cov(\epsilon_{i,0},\eta_i)+\frac{\sigma^2_{trend}}{\gamma})((1-\gamma)^t-1)
\end{aligned}$$

So,

$$Var(E'_{i,t}) = Var(\tau_{i,t}) + t^2\sigma^2_{trend} + 2t(Cov(\epsilon_{i,0},\eta_i)+\frac{\sigma^2_{trend}}{\gamma})((1-\gamma)^t-1) \tag{A55}$$

In the stationary regime, we have : $Var(\epsilon_{i,0}) = Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))(2-2(1-\gamma)^t)$ and $Cov(\epsilon_{i,0},\eta_i) = -\frac{\sigma^2_{trend}}{\gamma}$. Then, we obtain : $Var(E'_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))(2-2(1-\gamma)^t)+t^2\sigma^2_{trend}$

Covariance of residuals

We have

$$\begin{aligned}
Cov(E'_{i,t},E'_{j,t-s}) &= Cov(\tau_{i,t}+\eta_i t,\tau_{j,t-s}+\eta_j(t-s))\\
&= Cov(\tau_{i,t},\tau_{j,t-s}) + (t-s)Cov(\tau_{i,t},\eta_j) + tCov(\eta_i,\tau_{j,t-s}) + t(t-s)Cov(\eta_i,\eta_j)\\
&= Cov(\epsilon_{i,t},\epsilon_{j,t-s}) + Cov(\epsilon_{i,0},\epsilon_{j,0})(1-(1-\gamma)^t-(1-\gamma)^{t-s}) + (t-s)Cov(\tau_{i,t},\eta_j)\\
&+ tCov(\eta_i,\tau_{j,t-s}) + t(t-s)Cov(\eta_i,\eta_j)
\end{aligned} \tag{A56}$$

Case 1 : $\mathbf{i=j}$

If $i=j$, we have :

$$\begin{aligned}
Cov(E'_{i,t},E'_{i,t-s}) &= Cov(\epsilon_{i,t},\epsilon_{i,t-s}) + Var(\epsilon_{i,0})(1-(1-\gamma)^t-(1-\gamma)^{t-s}) + t(t-s)\sigma^2_{trend}\\
&+ (Cov(\epsilon_{i,0},\eta_i)+\frac{\sigma^2_{trend}}{\gamma})((t-s)((1-\gamma)^t-1)+t((1-\gamma)^{t-s}-1))
\end{aligned}$$

In the stationary regime, we have : $Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})\times\delta_{i,j})$, $Var(\epsilon_{i,0}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))$, $Cov(\epsilon_{i,0},\eta_i) = -\frac{\sigma^2_{trend}}{\gamma}$ and $Cov(\tau_{i,t},\eta_i)=0$.

So, we obtain :

$$\begin{aligned}
Cov(E'_{i,t},E'_{j,t-s}) &= \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})) + t(t-s)\sigma^2_{trend}\\
&+ \frac{1}{1-(1-\gamma)^2}(\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))(1-(1-\gamma)^t-(1-\gamma)^{t-s})
\end{aligned}$$

Case 2 : $\mathbf{i\neq j}$

If $i\neq j$, we have $Cov(\eta_i,\eta_j)=0$, $Cov(\eta_i,\tau_{j,t-s})=0$ and $Cov(\eta_j,\tau_{i,t})=0$. So, $Cov(E'_{i,t},E'_{j,t-s}) = Cov(\epsilon_{i,t},\epsilon_{j,t-s}) + Cov(\epsilon_{i,0},\epsilon_{j,0})(1-(1-\gamma)^t-(1-\gamma)^{t-s})$

In the stationary regime, we have : $Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2}$ and $Cov(\epsilon_{i,0},\epsilon_{j,0}) = \frac{1}{1-(1-\gamma)^2}(\rho^{d_{i,j}}\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))$. Then, we obtain :

$$Cov(E'_{i,t},E'_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} + \frac{1}{1-(1-\gamma)^2}(\rho^{d_{i,j}}\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma}))(1-(1-\gamma)^t-(1-\gamma)^{t-s})$$

Denoting $\mathbf{\Psi}$ the variance-covariance matrix of of the residuals $E'_{i,t}$, we can easily see that $\mathbf{\Psi}$ is a sum of matrices.

Stationary form
The stationary form of $\mathbf{\Psi}$ is :

$$\begin{aligned}\mathbf{\Psi} = & \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \dots & \mathbf{P}_{1,S}\\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \dots & \mathbf{P}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & \dots & \mathbf{P}_{S,S}\end{bmatrix} + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix}\mathbf{J} & 0 & \dots & 0\\ 0 & \mathbf{J} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{J}\end{bmatrix}\\ & + \sigma^2_{trend}\begin{bmatrix}\mathbf{C}_{1,1} & 0 & \dots & 0\\ 0 & \mathbf{C}_{2,2} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \mathbf{C}_{S,S}\end{bmatrix} + \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix}\mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,S}\\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & \dots & \mathbf{Q}_{S,S}\end{bmatrix}\\ & + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix}\mathbf{L}_{1,1} & \mathbf{L}_{1,2} & \dots & \mathbf{L}_{1,S}\\ \mathbf{L}_{2,1} & \mathbf{L}_{2,2} & \dots & \mathbf{L}_{2,S}\\ \dots & \dots & \dots & \dots\\ \mathbf{L}_{S,1} & \mathbf{L}_{S,2} & \dots & \mathbf{L}_{S,S}\end{bmatrix}\end{aligned} \tag{A57}$$

with

$$\mathbf{P}_{i,j} = \begin{bmatrix}\rho^{d_{i,j}}\sigma^2 & (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2\\ (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \rho^{d_{i,j}}\sigma^2 & \dots & \dots\\ \dots & \dots & \dots & \dots\\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 & \dots & \dots & \rho^{d_{i,j}}\sigma^2\end{bmatrix}$$

$$\mathbf{J} = \begin{bmatrix}1 & 1 & \dots & 1\\ 1 & 1 & \dots & \dots\\ \dots & \dots & \dots & \dots\\ 1 & \dots & \dots & 1\end{bmatrix}$$

$$\mathbf{C}_{i,j} = \begin{bmatrix}c^{0,0}_{i,j} & c^{0,\beta}_{i,j} & \dots & c^{0,\beta(T-1)}_{i,j}\\ c^{\beta,0}_{i,j} & c^{\beta,\beta}_{i,j} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ c^{\beta(T-1),0}_{i,j} & \dots & \dots & c^{\beta(T-1),\beta(T-1)}_{i,j}\end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix}\chi^{0,0}_{i,j} & \chi^{0,\beta}_{i,j} & \dots & \chi^{0,\beta(T-1)}_{i,j}\\ \chi^{\beta,0}_{i,j} & \chi^{\beta,\beta}_{i,j} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ \chi^{\beta(T-1),0}_{i,j} & \dots & \dots & \chi^{(\beta(T-1),\beta(T-1)}_{i,j}\end{bmatrix}$$

and

$$\mathbf{L}_{i,j} = \begin{bmatrix}l^{0,0}_{i,j} & l^{0,\beta}_{i,j} & \dots & l^{0,\beta(T-1)}_{i,j}\\ l^{\beta,0}_{i,j} & l^{\beta,\beta}_{i,j} & \dots & \dots\\ \dots & \dots & \dots & \dots\\ l^{\beta(T-1),0}_{i,j} & \dots & \dots & l^{\beta(T-1),\beta(T-1)}_{i,j}\end{bmatrix}$$

where $c^{t_1,t_2}_{i,j} = t_1 t_2$, $\chi^{t_1,t_2}_{i,j} = \rho^{d_{i,j}}(1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2})$ and $l^{t_1,t_2}_{i,j} = 1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2}$

Non stationary form
The non stationary form of $\mathbf{\Psi}$ is :

$$\mathbf{\Psi} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma_0^2\sigma_{trend}^2}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & ... & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & ... & \mathbf{Q}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & ... & \mathbf{Q}_{S,S} \end{bmatrix} + \sigma_{trend}^2\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix}$$
$$+ \frac{\sigma_{trend}^2}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{L}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{L}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{L}_{S,S} \end{bmatrix} + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma_{trend}^2(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}\begin{bmatrix} \mathbf{G}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{G}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{G}_{S,S} \end{bmatrix} \tag{A58}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,\beta} & \dots & \zeta_{i,j}^{0,\beta(T-1)} \\ \zeta_{i,j}^{\beta,0} & \zeta_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{\beta(T-1),0} & \dots & \dots & \zeta_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix} q_{i,j}^{0,0} & q_{i,j}^{0,\beta} & \dots & q_{i,j}^{0,\beta(T-1)} \\ q_{i,j}^{\beta,0} & q_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ q_{i,j}^{\beta(T-1),0} & \dots & \dots & q_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{C}_{i,j} = \begin{bmatrix} c_{i,j}^{0,0} & c_{i,j}^{0,\beta} & \dots & c_{i,j}^{0,\beta(T-1)} \\ c_{i,j}^{\beta,0} & c_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ c_{i,j}^{\beta(T-1),0} & \dots & \dots & c_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{L}_{i,j} = \begin{bmatrix} l_{i,j}^{0,0} & l_{i,j}^{0,\beta} & \dots & l_{i,j}^{0,\beta(T-1)} \\ l_{i,j}^{\beta,0} & l_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ l_{i,j}^{\beta(T-1),0} & \dots & \dots & l_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

and

$$\mathbf{G}_{i,j} = \begin{bmatrix} g_{i,j}^{0,0} & g_{i,j}^{0,\beta} & \dots & g_{i,j}^{0,\beta(T-1)} \\ g_{i,j}^{\beta,0} & g_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ g_{i,j}^{\beta(T-1),0} & \dots & \dots & g_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

with $\zeta_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(\sigma_0^2((1-\gamma)^{t_1+t_2}+1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2})+(1-\gamma)^{|t_1-t_2|}-(1-\gamma)^{t_1+t_2})$,
$q_{i,j}^{t_1,t_2} = (1-\gamma)^{t_1+t_2}+1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2}$, $l_{i,j}^{t_1,t_2} = 1-(1-\gamma)^{t_1+t_2-|t_1-t_2|}$,
$c_{i,j}^{t_1,t_2} = t_1t_2+\frac{t_2}{\gamma}((1-\gamma)^{t_1}-1)+\frac{t_1}{\gamma}((1-\gamma)^{t_2}-1)-\frac{(1-\gamma)^{t_1}-(1-\gamma)^{t_2}}{\gamma}\times\frac{1-(1-\gamma)^{t_2}}{\gamma}$ and
$g_{i,j}^{t_1,t_2} = t_2((1-\gamma)^{t_1}-1)+t_1((1-\gamma)^{t_2}-1)-\frac{1-(1-\gamma)^{t_2}}{\gamma}((1-\gamma)^{t_1}+(1-\gamma)^{t_2})-\frac{(1-\gamma)^{t_1+t_2-|t_1-t_2|}-(1-\gamma)^{t_1+t_2}}{\gamma}$

Model with observation error
If there is an observation error, the model (eq.A54) becomes :

$$\begin{aligned} Y_{i,t} &= ln\tilde{N}_{i,t} - ln\tilde{N}_{i,0} \\ &= lnN_{i,t} - lnN_{i,0} + v_{i,t} - v_{i,0} \\ &= rt + \eta_i t + \epsilon_{i,t} - \epsilon_{i,0} + v_{i,t} - v_{i,0} \\ &= -(\frac{r+0.5\sigma^2}{\gamma}+c) + rt + (1-\gamma)^t(c+\frac{r+0.5\sigma^2}{\gamma}) + E''_{i,t} \end{aligned} \tag{A59}$$

where $E''_{i,t} = E'_{i,t} + v_{i,t} - v_{i,0}$ is a normally distributed multivariate variable with unknown variance-covariance matrix and $Mean(E''_{i,t}) = 0$.
Let derive variance and covariance of $E''_{i,t}$.

Variance

$$Var(E''_{i,t}) = Var(E'_{i,t}) + Var(v_{i,t}) + Var(v_{i,0}) - 2Cov(v_{i,t}, v_{i,0})$$

Referring to eq.A43, we obtain

$$Var(E''_{i,t}) = Var(E'_{i,t}) + 2\sigma^2_{obs}\mathbb{I}_{t>0} \tag{A60}$$

In the stationary regime, we obtain $Var(E''_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))(2 - 2(1-\gamma)^t) + t^2\sigma^2_{trend} + 2\sigma^2_{obs}\mathbb{I}_{t>0}$.

Covariance

$$\begin{aligned} Cov(E''_{i,t}, E''_{j,t-s}) &= Cov(E'_{i,t} + v_{i,t} - v_{i,0}, E'_{j,t-s} + v_{j,t-s} - v_{j,0}) \\ &= Cov(E'_{i,t}, E'_{j,t-s}) + Cov(E'_{i,t}, v_{j,t-s}) - Cov(E'_{i,t}, v_{j,0}) + Cov(v_{i,t}, E'_{j,t-s}) + Cov(v_{i,t}, v_{j,t-s}) \\ &\quad - Cov(v_{i,t}, v_{j,0}) - Cov(v_{i,0}, E'_{j,t-s}) - Cov(v_{i,0}, v_{j,t-s}) + Cov(v_{i,0}, v_{j,0}) \end{aligned} \tag{A61}$$

Case 1 : $i \neq j$
If $i \neq j$, we have :

$$Cov(E''_{i,t}, E''_{j,t-s}) = Cov(E'_{i,t}, E'_{j,t-s})$$

because $Cov(v_{i,t}, v_{j,s}) = 0$ and $Cov(E'_{i,t}, v_{j,t}) = 0$

In the stationary regime, we obtain
$Cov(E''_{i,t}, E''_{j,t-s}) = \frac{(1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma)^2} + \frac{1}{1-(1-\gamma)^2}(\rho^{d_{i,j}}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))(1 - (1-\gamma)^t - (1-\gamma)^{t-s})$

Case 2 : $i = j$
If $i = j$, we have :

$$\begin{aligned} Cov(E''_{i,t}, E''_{i,t-s}) &= Cov(E'_{i,t}, E'_{i,t-s}) + Cov(E'_{i,t}, v_{i,t-s}) - Cov(E'_{i,t}, v_{i,0}) + Cov(v_{i,t}, E'_{i,t-s}) + Cov(v_{i,t}, v_{i,t-s}) \\ &\quad - Cov(v_{i,t}, v_{i,0}) - Cov(v_{i,0}, E'_{i,t-s}) - Cov(v_{i,0}, v_{i,t-s}) + \sigma^2_{obs} \end{aligned}$$

- If $s \neq 0$
  We have :

  $$Cov(E''_{i,t}, E''_{i,t-s}) = Cov(E'_{i,t}, E'_{i,t-s}) - Cov(v_{i,t}, v_{i,0}) - Cov(v_{i,0}, v_{i,t-s}) + \sigma^2_{obs}$$

  because $Cov(E'_{i,t}, v_{i,t-s}) = 0$ and $Cov(v_{i,t}, v_{i,t-s}) = 0$. Moreover, $Cov(v_{i,t}, v_{i,0}) = \sigma^2_{obs}\mathbb{I}_{t=0}$ and $Cov(v_{i,0}, v_{i,t-s}) = \sigma^2_{obs}\mathbb{I}_{t=s}$.
  So,

  $$Cov(E''_{i,t}, E''_{i,t-s}) = Cov(E'_{i,t}, E'_{i,t-s}) + \sigma^2_{obs}(1 - \mathbb{I}_{t=0} - \mathbb{I}_{t=s})$$

  In the stationary regime,

  $$\begin{aligned} Cov(E''_{i,t}, E''_{j,t-s}) = &\frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}) + t(t-s)\sigma^2_{trend} + \sigma^2_{obs}(1 - \mathbb{I}_{t=0} - \mathbb{I}_{t=s}) \\ &+ \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend}(1 + \frac{2-2\gamma}{\gamma}))(1 - (1-\gamma)^t - (1-\gamma)^{t-s}) \end{aligned}$$

- If $s = 0$
  We have :

  $$Cov(E''_{i,t}, E''_{i,t}) = Var(E''_{i,t})$$

Denoting $\boldsymbol{\Delta}$ the variance-covariance matrix of the residuals $E''_{i,t}$, we have

Stationary form

$$\boldsymbol{\Delta} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{J} & 0 & ... & 0 \\ 0 & \mathbf{J} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{J} \end{bmatrix} + \sigma^2_{trend}\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix}$$
$$+ \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & ... & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & ... & \mathbf{Q}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & ... & \mathbf{Q}_{S,S} \end{bmatrix} + \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{L}_{1,1} & \mathbf{L}_{1,2} & ... & \mathbf{L}_{1,S} \\ \mathbf{L}_{2,1} & \mathbf{L}_{2,2} & ... & \mathbf{L}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{L}_{S,1} & \mathbf{L}_{S,2} & ... & \mathbf{L}_{S,S} \end{bmatrix} + \sigma^2_{obs}\begin{bmatrix} \mathbf{K} & 0 & ... & 0 \\ 0 & \mathbf{K} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{K} \end{bmatrix} \tag{A62}$$

with

$$\mathbf{P}_{i,j} = \begin{bmatrix} \rho^{d_{i,j}}\sigma^2 & (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \dots & (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 \\ (1-\gamma)^{\beta}\rho^{d_{i,j}}\sigma^2 & \rho^{d_{i,j}}\sigma^2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1-\gamma)^{\beta(T-1)}\rho^{d_{i,j}}\sigma^2 & \dots & \dots & \rho^{d_{i,j}}\sigma^2 \end{bmatrix}$$

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 1 & \dots & \dots & 1 \end{bmatrix}$$

$$\mathbf{C}_{i,j} = \begin{bmatrix} c_{i,j}^{0,0} & c_{i,j}^{0,\beta} & \dots & c_{i,j}^{0,\beta(T-1)} \\ c_{i,j}^{\beta,0} & c_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ c_{i,j}^{\beta(T-1),0} & \dots & \dots & c_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix} \chi_{i,j}^{0,0} & \chi_{i,j}^{0,\beta} & \dots & \chi_{i,j}^{0,\beta(T-1)} \\ \chi_{i,j}^{\beta,0} & \chi_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \chi_{i,j}^{\beta(T-1),0} & \dots & \dots & \chi_{i,j}^{(\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{L}_{i,j} = \begin{bmatrix} l_{i,j}^{0,0} & l_{i,j}^{0,\beta} & \dots & l_{i,j}^{0,\beta(T-1)} \\ l_{i,j}^{\beta,0} & l_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ l_{i,j}^{\beta(T-1),0} & \dots & \dots & l_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

and

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 2 & 1 & 1 \\ \dots & 1 & 2 & \dots \\ 0 & 1 & \dots & 2 \end{bmatrix}$$

where $c_{i,j}^{t_1,t_2} = t_1 t_2$, $\chi_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2})$ and $l_{i,j}^{t_1,t_2} = 1-(1-\gamma)^{t_1}-(1-\gamma)^{t_2}$

Non stationary form

$$\boldsymbol{\Delta} = \frac{\sigma^2}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & ... & \mathbf{P}_{1,S} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & ... & \mathbf{P}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{P}_{S,1} & \mathbf{P}_{S,2} & ... & \mathbf{P}_{S,S} \end{bmatrix} + \frac{\sigma_0^2\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & ... & \mathbf{Q}_{1,S} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & ... & \mathbf{Q}_{2,S} \\ ... & ... & ... & ... \\ \mathbf{Q}_{S,1} & \mathbf{Q}_{S,2} & ... & \mathbf{Q}_{S,S} \end{bmatrix} + \sigma^2_{trend}\begin{bmatrix} \mathbf{C}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{C}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C}_{S,S} \end{bmatrix}$$
$$+ \frac{\sigma^2_{trend}}{1-(1-\gamma)^2}(1+\frac{2-2\gamma}{\gamma})\begin{bmatrix} \mathbf{L}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{L}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{L}_{S,S} \end{bmatrix} + \frac{\theta\sigma_0\sigma_{trend}\sqrt{\sigma^2+\sigma^2_{trend}(1+\frac{2-2\gamma}{\gamma})}}{\sqrt{1-(1-\gamma)^2}}\begin{bmatrix} \mathbf{G}_{1,1} & 0 & ... & 0 \\ 0 & \mathbf{G}_{2,2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{G}_{S,S} \end{bmatrix} + \sigma^2_{obs}\begin{bmatrix} \mathbf{K} & 0 & ... & 0 \\ 0 & \mathbf{K} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{K} \end{bmatrix} \tag{A63}$$

where

$$\mathbf{P}_{i,j} = \begin{bmatrix} \zeta_{i,j}^{0,0} & \zeta_{i,j}^{0,\beta} & \dots & \zeta_{i,j}^{0,\beta(T-1)} \\ \zeta_{i,j}^{\beta,0} & \zeta_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \zeta_{i,j}^{\beta(T-1),0} & \dots & \dots & \zeta_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{Q}_{i,j} = \begin{bmatrix} q_{i,j}^{0,0} & q_{i,j}^{0,\beta} & \dots & q_{i,j}^{0,\beta(T-1)} \\ q_{i,j}^{\beta,0} & q_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ q_{i,j}^{\beta(T-1),0} & \dots & \dots & q_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{C}_{i,j} = \begin{bmatrix} c_{i,j}^{0,0} & c_{i,j}^{0,\beta} & \dots & c_{i,j}^{0,\beta(T-1)} \\ c_{i,j}^{\beta,0} & c_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ c_{i,j}^{\beta(T-1),0} & \dots & \dots & c_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{L}_{i,j} = \begin{bmatrix} l_{i,j}^{0,0} & l_{i,j}^{0,\beta} & \dots & l_{i,j}^{0,\beta(T-1)} \\ l_{i,j}^{\beta,0} & l_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ l_{i,j}^{\beta(T-1),0} & \dots & \dots & l_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

$$\mathbf{G}_{i,j} = \begin{bmatrix} g_{i,j}^{0,0} & g_{i,j}^{0,\beta} & \dots & g_{i,j}^{0,\beta(T-1)} \\ g_{i,j}^{\beta,0} & g_{i,j}^{\beta,\beta} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ g_{i,j}^{\beta(T-1),0} & \dots & \dots & g_{i,j}^{\beta(T-1),\beta(T-1)} \end{bmatrix}$$

and

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 2 & 1 & 1 \\ \dots & 1 & 2 & \dots \\ 0 & 1 & \dots & 2 \end{bmatrix}$$

with $\zeta_{i,j}^{t_1,t_2} = \rho^{d_{i,j}}(\sigma_0^2((1-\gamma)^{t_1+t_2} + 1 - (1-\gamma)^{t_1} - (1-\gamma)^{t_2}) + (1-\gamma)^{|t_1-t_2|} - (1-\gamma)^{t_1+t_2})$,
$q_{i,j}^{t_1,t_2} = (1-\gamma)^{t_1+t_2} + 1 - (1-\gamma)^{t_1} - (1-\gamma)^{t_2}$, $l_{i,j}^{t_1,t_2} = 1 - (1-\gamma)^{t_1+t_2-|t_1-t_2|}$,
$c_{i,j}^{t_1,t_2} = t_1 t_2 + \frac{t_2}{\gamma}((1-\gamma)^{t_1} - 1) + \frac{t_1}{\gamma}((1-\gamma)^{t_2} - 1) - \frac{(1-\gamma)^{t_1} - (1-\gamma)^{t_2}}{\gamma} \times \frac{1-(1-\gamma)^{t_2}}{\gamma}$ and
$g_{i,j}^{t_1,t_2} = t_2((1-\gamma)^{t_1} - 1) + t_1((1-\gamma)^{t_2} - 1) - \frac{1-(1-\gamma)^{t_2}}{\gamma}((1-\gamma)^{t_1} + (1-\gamma)^{t_2}) - \frac{(1-\gamma)^{t_1+t_2-|t_1-t_2|} - (1-\gamma)^{t_1+t_2}}{\gamma}$

Conditional appraoch
Here the term $\eta_i t$ will be deterministic and then it will be contained into the mean of the model not into the residuals. The regression model (eq.A47) becomes :

$$lnN_{i,t} = \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + (1-\gamma)^t (c + \frac{r + 0.5\sigma^2}{\gamma}) + \eta_i t + \epsilon'_{i,t} \tag{A64}$$

The residuals is $\epsilon'_{i,t}$. In the stationary regime, we obtain

$$lnN_{i,t} = \mu + \omega x_i - \frac{r + 0.5\sigma^2}{\gamma} + rt + \eta_i t + \epsilon'_{i,t}$$

Now we can derive mean and variance-correlation structure of the residuals, $\epsilon'_{i,t}$, for this model. Since $Mean(\epsilon'_{i,t}) = 0$, $Var(\epsilon'_{i,t}) = Var(\epsilon_{i,t})$ and $Cov(\epsilon'_{i,t}, \epsilon'_{j,t-s}) = Cov(\epsilon_{i,t}, \epsilon_{j,t-s})$, then the variance- covariance matrix of the residuals $\epsilon'_{i,t}$ is the matrix $\mathbf{\Pi}$ of the DFLE (see eq.A16 and A17).

Model with observation error

Like in marginal model, one may want to introduce an additional observation error term in this statistical model by adding the diagonal error matrix below called z to $\mathbf{\Pi}$.

$$z = \frac{\sigma^2_{obs}}{1-(1-\gamma)^2}\begin{bmatrix} \mathbf{I}_T & 0 & \dots & 0 \\ 0 & \mathbf{I}_T & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{I}_T \end{bmatrix}$$

Introduction of strata

Let suppose that the population is divided into g strata with different trend parameters. That means that each sub-population i of each strata has a different deterministic trend r_i, and that these trends are normally and independently distributed with mean r_g and variance $\sigma^2_{trend,g}$. In this case, the deviations from local equilibrium (DFLE) becomes

$$\epsilon_{i,t} = (1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2 \tag{A65}$$

where η_i is an independent and normally distributed random variable with mean 0 and variance $\sigma^2_{trend,g(i)}$.

Mean of DFLE

Based on eq.A11, we obtain

$$Mean(\epsilon_{i,t}) = (1-\gamma)^t Mean(\epsilon_{i,0}) - (r_{g(i)} + 0.5\sigma^2)\frac{1-(1-\gamma)^t}{\gamma} \tag{A66}$$

If 0<γ<1, the sequence $Mean(\epsilon_{i,t})$ converges to $\frac{-(r_{g(i)}+0.5\sigma^2)}{\gamma}$ when t increases to infinity.

Variance

$$\begin{aligned} Var(\epsilon_{i,t}) &= Var((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2) \\ &= (1-\gamma)^2 Var(\epsilon_{i,t-1}) + \sigma^2 + \sigma^2_{trend,g(i)} - 2(1-\gamma)Cov(\epsilon_{i,t-1}, \eta_i) \end{aligned}$$

because $u_{i,t}$ and η_i are independent. Referring to eq.A13, we have

$$Var(\epsilon_{i,t}) = (1-\gamma)^{2t}Var(\epsilon_{i,0}) + (\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma}{\gamma}))\frac{1-(1-\gamma)^{2t}}{1-(1-\gamma)^2} - 2(1-\gamma)^t(Cov(\epsilon_{i,0},\eta_i) + \frac{\sigma^2_{trend,g(i)}}{\gamma})\frac{1-(1-\gamma)^t}{\gamma}) \tag{A67}$$

Asymptotically in t, we obtain $Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma)^2}(\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma}{\gamma}))$ because $(1-\gamma)^t \to 0$.

Covariance

We have

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) &= Cov((1-\gamma)\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2, (1-\gamma)\epsilon_{j,t-s-1} + u_{j,t-s} - r_{g(j)} - \eta_j - 0.5\sigma^2) \\ &= (1-\gamma)^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma)Cov(\epsilon_{i,t-1}, u_{j,t-s}) - (1-\gamma)Cov(\epsilon_{i,t-1}, \eta_j) \\ &+ (1-\gamma)Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) - (1-\gamma)Cov(\eta_i, \epsilon_{j,t-s-1}) + Cov(\eta_i, \eta_j) \end{aligned}$$

Based on eq.A15, asymptotically in t, with fixed s, we have :

$$Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \frac{1}{1-(1-\gamma)^2}((1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma}{\gamma}) \times \delta_{i,j})$$

Whereas the non-asymptotic expression of covariance (see eq.A21) is :

$$Cov(\epsilon_{i,t},\epsilon_{j,t-s}) = \begin{cases} (1-\gamma)^{2t-s}Cov(\epsilon_{i,0},\epsilon_{j,0}) + (1-\gamma)^{|s|}\rho^{d_{i,j}}\sigma^2\frac{1-(1-\gamma)^{2(t-s)}}{1-(1-\gamma)^2} & \text{if } i \neq j \\ (1-\gamma)^{2min(t,t-s)}Var(\epsilon_{i,0}) - (1-\gamma)^{2min(t,t-s)}Cov(\epsilon_{i,0},\eta_i)\frac{1-(1-\gamma)^{|s|}}{\gamma} + A_1\frac{1-(1-\gamma)^{2min(t,t-s)}}{1-(1-\gamma)^2} & \\ -A_2(1-\gamma)^t\frac{1-(1-\gamma)^{min(t,t-s)}}{\gamma} & \text{if } i = j \end{cases} \tag{A68}$$

where $A_1 = (1-\gamma)^{|s|}\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma}{\gamma})$ and $A_2 = (Cov(\epsilon_{i,0},\eta_i) + \frac{\sigma^2_{trend,g(i)}}{\gamma})(1+(1-\gamma)^{-s})$

Variation of density-dependance

Let suppose that in each strata the density-dependance is different than the other. That means that each sub-population i of each strata is submitted to a different density-dependence, with γ_g, the strength (intensity) of that dependence in the strata. ρ and σ are assumed to be constant from one strata to another. In this case, the DFLE (eq.A65) becomes :

$$\epsilon_{i,t} = (1-\gamma_{g(i)})\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2 \tag{A69}$$

Mean of DFLE
Based on eq.A66, we have

$$Mean(\epsilon_{i,t}) = (1-\gamma_{g(i)})^t Mean(\epsilon_{i,0}) - (r_{g(i)} + 0.5\sigma^2)\frac{1-(1-\gamma_{g(i)})^t}{\gamma_{g(i)}} \tag{A70}$$

If $0<\gamma_{g(i)}<1$, the sequence $Mean(\epsilon_{i,t})$ converges to $\frac{-(r_{g(i)}+0.5\sigma^2)}{\gamma_{g(i)}}$ when t increases to infinity.

Variance
From eq.A67, we have

$$\begin{aligned} Var(\epsilon_{i,t}) &= (1-\gamma_{g(i)})^{2t} Var(\epsilon_{i,0}) + (\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma_{g(i)}}{\gamma_{g(i)}}))\frac{1-(1-\gamma_{g(i)})^{2t}}{1-(1-\gamma_{g(i)})^2} \\ &- 2(1-\gamma_{g(i)})^t(Cov(\epsilon_{i,0},\eta_i) + \frac{\sigma^2_{trend,g(i)}}{\gamma_{g(i)}})\frac{1-(1-\gamma_{g(i)})^t}{\gamma_{g(i)}}) \end{aligned} \tag{A71}$$

Asymptotically in t, we obtain $Var(\epsilon_{i,t}) = \frac{1}{1-(1-\gamma_{g(i)})^2}(\sigma^2+\sigma^2_{trend,g(i)}(1+\frac{2-2\gamma_{g(i)}}{\gamma_{g(i)}}))$ because $(1-\gamma_{g(i)})^t \to 0$.

Covariance
We have

$$\begin{aligned} Cov(\epsilon_{i,t},\epsilon_{j,t-s}) &= Cov((1-\gamma_{g(i)})\epsilon_{i,t-1} + u_{i,t} - r_{g(i)} - \eta_i - 0.5\sigma^2, (1-\gamma_{g(j)})\epsilon_{j,t-s-1} + u_{j,t-s} - r_{g(j)} - \eta_j - 0.5\sigma^2) \\ &= (1-\gamma_{g(i)})(1-\gamma_{g(j)})Cov(\epsilon_{i,t-1},\epsilon_{j,t-s-1}) + (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1},u_{j,t-s}) - (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1},\eta_j) \\ &+ (1-\gamma_{g(j)})Cov(u_{i,t},\epsilon_{j,t-s-1}) + Cov(u_{i,t},u_{j,t-s}) - (1-\gamma_{g(j)})Cov(\eta_i,\epsilon_{j,t-s-1}) + Cov(\eta_i,\eta_j) \end{aligned} \tag{A72}$$

If i and j are in the same strata, $Cov(\epsilon_{i,t-1},u_{i,t-s}) = (1-\gamma_{g(i)})^{s-1}\rho^{d_{i,j}}\sigma^2$. If not, $Cov(\epsilon_{i,t-1},u_{i,t-s}) = 0$.

Case 1 : i = j
$i=j$ means that we are in the same strata and $\gamma_{g(i)} = \gamma_{g(j)}$. Then, we have:

$$\begin{aligned} Cov(\epsilon_{i,t},\epsilon_{j,t-s}) &= (1-\gamma_{g(i)})^2 Cov(\epsilon_{i,t-1},\epsilon_{i,t-s-1}) + (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1},u_{i,t-s}) - (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1},\eta_i) \\ &+ (1-\gamma_{g(i)})Cov(u_{i,t},\epsilon_{i,t-s-1}) + Cov(u_{i,t},u_{i,t-s}) - (1-\gamma_{g(i)})Cov(\eta_i,\epsilon_{i,t-s-1}) + \sigma^2_{trend,g(i)} \end{aligned}$$

Since $Cov(\epsilon_{i,t-1},\eta_i) = (1-\gamma_{g(i)})^{t-1}Cov(\epsilon_{i,0},\eta_i) - \sigma^2_{trend,g(i)}\frac{1-(1-\gamma_{g(i)})^{t-1}}{\gamma_{g(i)}}$ based on eq.A12, then, we obtain from eq.A17 :

$$\begin{aligned} Cov(\epsilon_{i,t},\epsilon_{j,t-s}) &= (1-\gamma_{g(i)})^{2min(t,t-s)}Var(\epsilon_{i,0}) - (1-\gamma_{g(i)})^{2min(t,t-s)}Cov(\epsilon_{i,0},\eta_i)\frac{1-(1-\gamma_{g(i)})^{|s|}}{\gamma_{g(i)}} \\ &+ A_1\frac{1-(1-\gamma_{g(i)})^{2min(t,t-s)}}{1-(1-\gamma_{g(i)})^2} - A_2(1-\gamma_{g(i)})^t\frac{1-(1-\gamma_{g(i)})^{min(t,t-s)}}{\gamma_{g(i)}} \end{aligned}$$

with $A_1 = (1-\gamma_{g(i)})^{|s|}\sigma^2 + \sigma^2_{trend,g(i)}(1+\frac{2-2\gamma_{g(i)}}{\gamma_{g(i)}})$ and $A_2 = (Cov(\epsilon_{i,0},\eta_i) + \frac{\sigma^2_{trend,g(i)}}{\gamma_{g(i)}})(1+(1-\gamma_{g(i)})^{-s})$

Asymptotically in t, we have : $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \rightarrow \frac{1}{1-(1-\gamma_{g(i)})^2}((1-\gamma_{g(i)})^{|s|}\sigma^2 + \sigma^2_{trend}(1+\frac{2-2\gamma_{g(i)}}{\gamma_{g(i)}}))$

Case 2 : $\mathbf{i \neq j}$

If $i \neq j$, $Cov(\eta_i, \eta_j) = 0$, $Cov(\epsilon_{i,t-1}, \eta_j) = 0$ and $Cov(\eta_i, \epsilon_{j,t-s-1}) = 0$. Then,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = 1 - \gamma_{g(i)})(1-\gamma_{g(j)})Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ + (1-\gamma_{g(j)})Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) \end{aligned}$$

Here, we have two sub-cases.

- *i and j are in the same stratum*
  In this case, we have $\gamma_{g(i)} = \gamma_{g(j)}$. And then,

$$\begin{aligned} Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma_{g(i)})^2 Cov(\epsilon_{i,t-1}, \epsilon_{j,t-s-1}) + (1-\gamma_{g(i)})Cov(\epsilon_{i,t-1}, u_{j,t-s}) \\ + (1-\gamma_{g(i)})Cov(u_{i,t}, \epsilon_{j,t-s-1}) + Cov(u_{i,t}, u_{j,t-s}) \end{aligned}$$

Based on the previous results, we have

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma_{g(i)})^{2t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma_{g(i)})^{|s|}\rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma_{g(i)})^{2min(t,t-s)}}{1-(1-\gamma_{g(i)})^2}$$

Asymptotically in t, we have $Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) \rightarrow \frac{(1-\gamma_{g(i)})^{|s|}\rho^{d_{i,j}}\sigma^2}{1-(1-\gamma_{g(i)})^2}$

- *i and j are in the different strata*
  Here, we have : $Cov(u_{i,t}, \epsilon_{j,t-s-1}) = 0$ and $Cov(u_{i,t}, u_{j,t-s}) = \begin{cases} 0 & \text{if } s \neq 0 \\ \rho^{d_{i,j}}\sigma^2 & \text{if } s = 0 \end{cases}$
  So, as we proved above, we have :

$$Cov(\epsilon_{i,t}, \epsilon_{j,t-s}) = (1-\gamma_{g(i)})^t(1-\gamma_{g(j)})^{t-s} Cov(\epsilon_{i,0}, \epsilon_{j,0}) + (1-\gamma_{g(i)})^{|s|}\rho^{d_{i,j}}\sigma^2 \frac{1-(1-\gamma_{g(i)})^{min(t,t-s)}(1-\gamma_{g(j)})^{min(t,t-s)}}{1-(1-\gamma_{g(i)})(1-\gamma_{g(j)}}$$

7.2. Appendix B : Calculus of simulation

Giving the following parameters :

- B total number of observations
- T the average number of visits per site
- $tmax$ maximum monitoring time
- T the average number of visits per site
- p the proportion of non-permanents sites
- S Total number of sites
- S_p Total number of permanents sites
- S_{np} Total number of non-permanents sites
- T_p number of visits per permanents sites
- T_{np} number of visits per non-permanents sites

Questions
The two main questions are :

- How many sites S do we need to visit in general ?
- How many time T_p do we need to visit permanents sites ?

Expression of T_p as a function of p and T
Let define S_p as a total number of permanents sites and S_{np} a total number of non-permanents sites.

We have :

$$T = \frac{S_p}{S} \times T_p + \frac{S_{np}}{S} \times 1$$

or

$$T = (1 - p)T_p + p$$

Since that we can express T_p as function of T:

$$T_p = \frac{T - p}{1 - p}$$

Expression of S
By definition we have $B = ST$:

$$S = \frac{B}{T}$$

Then we can deduce S_p and S_{np} :

$$S_p = \frac{B}{T}(1 - p)$$

$$S_{np} = \frac{B}{T}p$$

7.3. Appendix C : Steps of simulations

- **Step 1 :** Definition of monitoring scenario
- **Step 2 :** Choice of a monitoring strategy for the temporal distribution of visits
- **Step 3 :** Implementation of the variance-covariance matrix of residuals $\mathbf{\Phi}$ under stationary assumption
- **Step 4 :** Calculus of $Var(\hat{r})$ (variation of the estimator of log-linear temporal trend) and standard error $se(\hat{r})$ knowing $\mathbf{\Phi}$ for each scenario
- **Step 5 :** Reproduction of the plots

8. References

Barbier, S., R., Chevalier, Philippe Loussot, Laurent Berge, F., Gosselin. 2008. *Improving biodiversity indicators of sustainable forest management: Tree genus abundance rather than tree genus richness and dominance for understory vegetation in French lowland oak hornbeam forests.* Forest Ecology and Management 258S, S176–S186.

Carvalho, Silvia b., J., Gonçalves, A., Guisan and J. P., Honrado. 2016. *"MODEL-ASSISTED MONITORING OF BIODIVERSITY- Systematic site selection for multispecies monitoring networks"* Journal of Applied Ecology 53(5): 1305–1316. https://doi.org/10.1111/1365-2664.12505

de Gruijter, J., D., Brus, M., Bierkens, M., Knotters. 2006. *Sampling for natural resource monitoring.* Springer, Inc., Berlin.

Dennis, B., J., M., Ponciano, S., R., Lele, M., L., Taper and D., F., Staples. 2006. *Estimating density dependence, process noise, and observation error.* Ecol. Monogr. 76: 323-341.

Dixon, P., M. and J., H., K., Pechmann. 2005. *A statistical test to show negligible trend.* Ecology, 86(7), pp.1751–1756.

Faraway, J. J. 2006. *Extending the linear model with R - Generalized linear, Mixed effets and Nonparametric regression models.* Chapman & Hall/CRC.

Gosselin, F., Dallari, R., 2007. *Gestion forestièr et biodiversité.* ARB, île de France, 17 octobre 2019.

Gosselin, F., Gosselin, M., Paillet, Y., 2012. *Suivre l'état de la biodiversité forestière : pourquoi ? comment ?*, Revue Forestière Française 64, 683-700.

Hanski, I., M., Kuussaari and M., Nieminen. 1994. *Metapopulation Structure and Migration in the Butterfly Melitaea Cinxia* Ecology 75(3): 747-762. https://doi.org/10.2307/1941732

Kariya, T. and Kurata, H. 2004. *Generalized least squares.* John Wiley & Sons, England.

Lindenmayer, D.B., Likens, G.E., 2010. Improving ecological monitoring. Trends in Ecology and Evolution 25, 200-201.

McDonald, T.L., 2003. *Review of environmental monitoring methods: Survey designs.* Environmental Monitoring and Assessment 85, 277-292.

Nielsen, S.E., Haughland, D.L., Bayne, E., Schieck, J., 2009. *Capacity of large-scale, long-term biodiversity monitoring programmes to detect trends in species prevalence.* Biodiversity and Conservation 18, 2961-2978.

Pinheiro, J. C, and Bates, D. M. 2000. *Mixed-Effects Models in S and Vols*, Springer, Inc., New York, NY.

Rao, C. R, and Toutenburg, H. 1995. *Linear Models-Least Squares and Alternatives.* Springer, Inc., New York, NY.

UICN. 2012. *Catégories et Critères de la Liste rouge de l'UICN : Version 3.1.* Deuxième édition. Gland, Suisse et Cambridge, Royaume-Uni : UICN. vi + 32pp. Originalement publié en tant que IUCN Red List Categories and Criteria: Version 3.1. Second edition. (Gland, Switzerland and Cambridge, UK: IUCN, 2012).

UMS PatriNat, online. – The monitoring of the terrestrial biodiversity, https://www.patrinat.fr/fr/surveillance-de-la-biodiversite-terrestre-6806

Urquhart, N. S. and Kincaid, T. M. 1999. *Designs for detecting trend from repeated surveys of ecological resources.*-J. Agric. Biol. Environ. Stat. 4: 404-414.

Vos, P., Meelis, E., Ter Keurs, W.J., 2000. *A framework for the design of ecological monitoring programs as a tool for environmental and nature management.* Environmental Monitoring and Assessment 61, 317-344.

Wikle, C. K., Zammit-Mangion, A. and Cressie, N. 2019. *Spatio-temporal statistics with R.* Chapman & Hall/CRC.

Gore, Sharad D. 2008. *Sampling for Natural Resource Monitoring*. Vol. 103. 482. https://doi.org/10.1198/jasa.2008.s240.

KK, T. G. Gregoire, D. R. Brillinger, P. J. E. Russek-Cohen, W. G. Warren, and R. D. Wolfinger. 1998. *Modelling Longitudinal and Spatially Correlated Data: Methods, Applications, and Future Directions*. Vol. 93. 441. https://doi.org/10.2307/2669661.

Little, Roderick J. 2004. "To model or not to model? Competing modes of inference for finite population sampling." *Journal of the American Statistical Association* 99 (466): 546–56. https://doi.org/10.1198/016214504000000467.

Rhodes, Jonathan R., and Niclas Jonzén. 2011. "Monitoring temporal trends in spatially structured populations: how should sampling effort be allocated between space and time?" *Ecography* 34 (6): 1040–48. https://doi.org/10.1111/j.1600-0587.2011.06370.x.

Wikle, Christopher K., and J. Andrew Royle. 1999. "Space-time dynamic design of environmental monitoring networks." *Journal of Agricultural, Biological, and Environmental Statistics* 4 (4): 489–507. https://doi.org/10.2307/1400504.