A stage-structured hierarchical Bayes model for the babassu palm tree population dynamics – Estimated from anthropogenic open area data sets

Nikolay Sirakov, Bénédicte Fontez, Thérèse Libourel Rouge, Alessio dos Santos, Danielle Mitja, Patrice Loisel

A Stage-Structured Hierarchical Bayes Model for the Babassu Palm Tree Population Dynamics - Estimated from anthropogenic open area data sets

Nikolay Sirakov[a], Bénédicte Fontez[a,*], Thérèse Libourel[b], Alessio dos Santos[c], Danielle Mitja[b], Patrice Loisel[a]

[a] *MISTEA, INRA, Montpellier SupAgro, Univ Montpellier, Montpellier, France*
[b] *Espace-Dev, IRD, Univ Montpellier, Montpellier, France*
[c] *ISARH - Instituto Ambiental e de Recursos Hídricos*
UFRA - Universidade Federal Rural da Amazônia, Belém, Brasil

Abstract

The babassu palm tree (*Attalea speciosa* Mart. ex Spreng.) is an endemic species of Amazon forests, and has social and economical impact/utility. Deforestation highlights this palm tree in anthropogenic open areas (pastures and cultivated fields). Simultaneously, knowledge concerning the sustainable functioning of the species within these manmade environments is sorely lacking: its life cycle is not well known, and its population dynamics remains unstudied. In this study, our objective was to generate a model of the population dynamics of the babassu palm tree, validated by *in situ* analysis, to understand how babassu, a forest species, adapts to pastureland and, under certain conditions, becomes invasive. We propose a random matrix model with aggregated variables based on the biological stages of the species as the input. The probabilities of the between-stage transition matrix were modelled using a Dirichlet-multinomial model with a hierarchy taking geographical organization, i.e. transect level, into account. The integration of prior information was formulated through a Bayesian approach. This Bayesian hierarchical matrix model enabled us to demonstrate a bottleneck in the population dynamics and a high year-dependent mortality rate at an early stage.

*Corresponding author
Email address: benedicte.fontez@supagro.fr (Bénédicte Fontez)

Keywords: mathematical modelling, Bayesian statistics, population dynamics, babassu palm tree

---

Contents

1. Introduction

The primary threat to primary forests, long before climate change, is human action, Jha (2010). The majority of plant species that grow in tropical rainforests do not resist land use change at the time of cultivation, Fearnside (2005); Barlow et al. (2016). Some other tree and palm species are resilient, Anthelme et al. (2011); Norden et al. (2009). The babassu palm tree (*Attalea speciosa* Mart. ex Spreng.) which can reach 30 m height, is endemic to the South American Amazonian forest. It is sometimes present in agrosystems installed following clearing of the primary forest, while many other woody species have disappeared, Mitja et al. (2018); dos Santos and Mitja (2011). It persists years later in secondary forests resulting from the abandonment of crops and pastures, and it can even become invasive and hinder the development of cultivated species. While a portion of the population wants to eliminate it, another part wants to maintain it because of its multiple uses, Araújo and Lopes (2012).

Some plot-scale work has demonstrated higher densities of adult babassu in pastures than in the original forest, Anderson (1983); Campos et al. (2017), while other work has shown the opposite, Barot et al. (2005). A territory-wide study of the PA-Benfica smallholder farming community revealed this variability in the density of babassu cauliflower palms in primary forests as well as in pastures and secondary forests. In these environments, the few plots with very high densities are only observed in pastures or secondary forests, Mitja et al. (2018). In contrast, because the number of plots without cauliflower babassu is higher in pastures, the average babassu density in the study area is lower in pastures than in secondary or primary forests, Mitja et al. (2018). Thus, babassu has the potential to invade agrosystems even if this does not occur systematically. In fact, there is a high variability of the palm density in pastures potentially that is related to the initial density in the forest, depending on the cultivation practices of farmers (who can conserve or eliminate palm trees), Mitja et al. (2018). In pastures, babassu are influenced by factors such as sun exposure, Mitja and Ferraz (2001); Anderson et al. (1991), burning, chemical

or manual cleaning of plots, competition with forage grasses, burial of fruit in the soil and clear-cutting. These factors sometimes act in a contradictory way on fruits, seedlings and young individuals and adults and can counteract the high resilience of this species at certain stages, Mitja et al. (2018). In our study context, the studied pastures did not undergo significant anthropogenic actions such as chemical cleaning or clear-cutting.

To explain this ability of babassu to remain on pasture after forest clearing, the possibility of natural biological diversity within babassus (diversity of genetic, epigenetic origins or more generally due to genetic interaction with the environment) cannot be excluded, Bräutigam et al. (2013); Richards et al. (2017), although we do not currently have data to support these assumptions.

Previous works, focusing on all babassu individuals from seedlings to adults, have examined the life cycle divided into ontogenic stages: three stages for Campos et al. (2017); four stages for Anderson (1983); five stages for Barot et al. (2005); six stages for dos Santos et al. (2017). In the latter case, which concerns our field of study, the life cycle was divided into a seedling stage, four stages of young individuals and an adult stage. Stages 1, 2 and 3 are characterized by a rosette configuration and a buried terminal bud. Stage 4 corresponds to the presence of the terminal bud on the soil surface, protected by the leaf sheaths. Stage 5 (last young stage) corresponds to individuals with a terminal bud above ground that is very often still protected by the leaf sheath. Stage 6 is the adult stage, most often ascertained by a free stipe. Considering the stages, some authors have studied the population structure using a single "one shot" measure, Barot et al. (2005); Campos et al. (2017). Barot et al. (2005) obtained an inverted J-shaped curve for the primary tropical forest, characteristics of natural unharvested populations, Hall and Bawa (1993), while in pastures this structure was found to be different with a lower seedling density. These latter results have not been verified by studies by Campos et al. (2017) and Anderson (1983), who showed an "inverted J" curve not only for primary forest but also for pastures. However, the accepted interpretation that species with inverted J population structures are in equilibrium, while those with a small number of

young stages would correspond to declining populations, has been questioned. Indeed, the survival of the plant population would depend more on the growth rate of individuals in the various stages, Condit et al. (1998), which, if rapid, can maintain a sufficient density of adults. To determine whether the babassu will persist in this newly set-up farmland in the long term, it will be necessary to assess the growth rate and thus monitor the evolution of the population over time, Johnson et al. (1994) by studying the population dynamics through periodic measurements over the long term.

Recently, individual-based models have received great interest and have been used to predict population dynamics and the effects of management strategies Buckley et al. (2003). Following the ideas of Uchmaski and Grimm (1996) and Lomnicki (1999), the interest in individual-based models is to take the individual interactions (competition) and the resource dynamics explicitly into account. In our case, pastureland is a stable and non limited resource environment where competition between trees is negligible. An individual-based model could be interesting to ascertain the time spent at each stage of development and to define trajectories, but as the data are from a three-year experiment, the results may not be sufficiently extensive to be reliable. Because of the very low transition rates often observed between some stages, a longer observation period would be necessary to be of use. The population dynamics models appear to offer a simpler and better adapted framework for our study Zuidema and Zagt (2000). Two components must be taken into account when modelling: the ecological process and the observation process. In our context, the ecological process describes variation in abundance over transects and time (two periods of measurement). Demographic models represent a standard approach for abundance data that provides insight into population growth Caswell (1989). Deterministic matrix models have been used for trees in the following articles: Bullock (1980), Pinero et al. (1984), Enright and Watson (1992), Pinard (1993), Olmsted and Alvarez-Buylla (1995), Barot et al. (2000), Holm et al. (2008), and also Guarín and del Valle (2014). Demographic models include a deterministic part but stochastic elements can be added to accommodate any sources of variability (transects, time,

individuals) that affect the observations. To take into account the geographical structure of our data, palm trees are observed on different transects; thus defining the hierarchical model. This approach was advocated by Clark (2003) who insists that variability associated with space, time, and among individuals who are not accommodated by demographic models can make parameter estimates and growth rate predictions unrealistic. Stochastic models can be defined in different ways. In the method defined by Tuljapurkar and Caswell (1997), the stochastic matrix involves elasticity and stochasticity of the vital rate. In our case, we did not focus on one parameter only but on a global approach that allows one to deal with broader issues and topics from different fields (economy, social impact, biology, technical aspect of the management and so on). IPM (integrated population models) is another possible approach, which is also known as hierarchical or state-space models Eacker et al. (2017), Schaub and Abadi (2011). More generally than IPM, the Bayesian hierarchical model is well adapted and sufficiently flexible to incorporate hierarchical data and expert knowledge, Hobbs and Hooten (2015). It can be adjusted over successive years of observation and updated over time, posterior estimation is used from the first year as prior information for the second year estimation, and so on, as suggested by Parent and Rivot (2012).

To understand how babassu, which is a forest species, adapts to grazing and under certain conditions becomes invasive, this study proposes to develop a population dynamics model that does not currently exist for this species. This model will be a hierarchical stochastic matrix population model in the Bayesian framework. Beyond the knowledge produced, this model can be used in future species management projects to make forecasts.

This article covers the following topics:

1. "Materials" provides a description of the biology of the babassu, as well as a description of the study site and of the sampling.
2. "Model" provides a comprehensive description of all aspects of the modelling (matrix model, hierarchy and Bayesian framework).

3. "Results-Discussion" provides comments on the model validation and on population dynamics and the existence of a bottleneck. These sections are followed by a global conclusion.

2. Materials

2.1. Biology of babassu in pastures

The babassu life cycle can be organized in biological stages of development. Six stages have been defined varying from the seedling to the adult palm tree (Fig. 1, see Mitja et al. (2019) for more details). In addition, the reproductive stage and fructification of the babassu has been studied by Barot et al. (2005) and Anderson et al. (1991) providing knowledge for understanding the reproductive process and recruitment in our modelling.

2.2. Study site

Our data was collected over three years of field monitoring (from 2013 to 2015). This monitoring was carried out on the Benfica Agriculture Installation Project (*Projeto de Assentamento*) site which is located in the Brazilian State of Pará (Fig. 2).

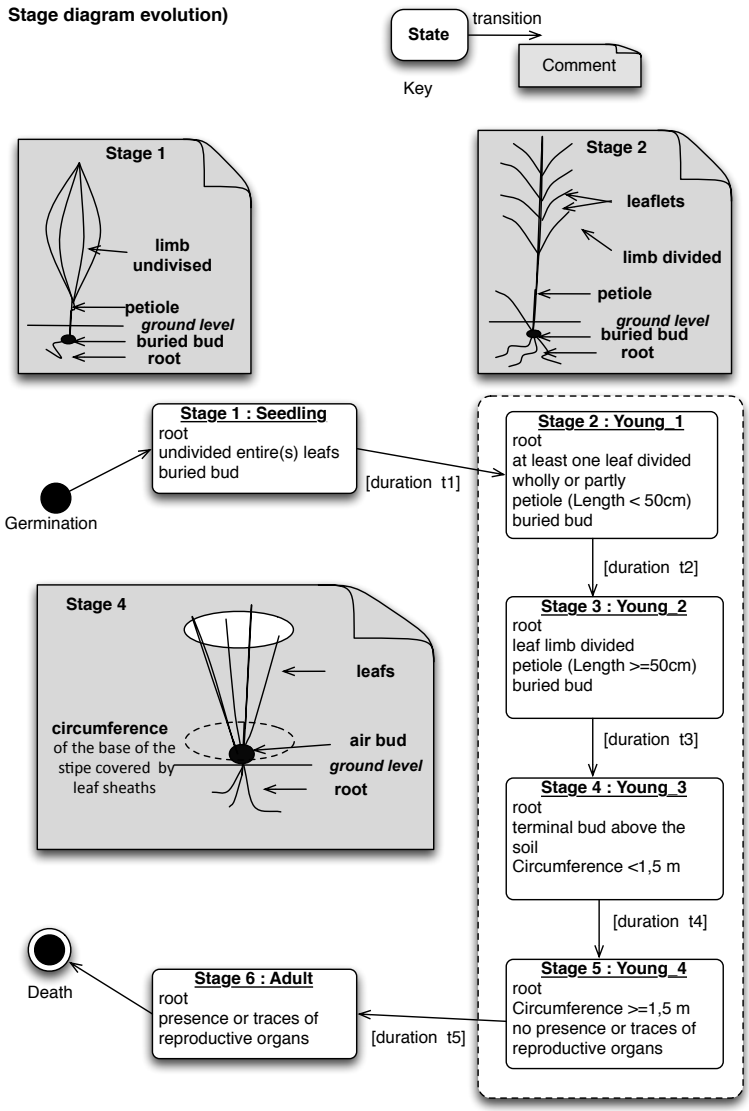
This territory is located within the Amazonian rain forest on the territory of the Itupiranga municipality and covers 10,026 ha. The shape of Benfica is irregular and lies within a quadrilateral space formed by the straight line connection at the coordinates 05°12'20" south latitude, 05°20'40" south latitude and 49°48'00" west longitude, 49°56'40" west longitude.

The particularity of this site is that it includes man-made spaces that, since 1989, have been progressively colonized by one thousand people, Sampaio (2008).

2.3. Field sampling

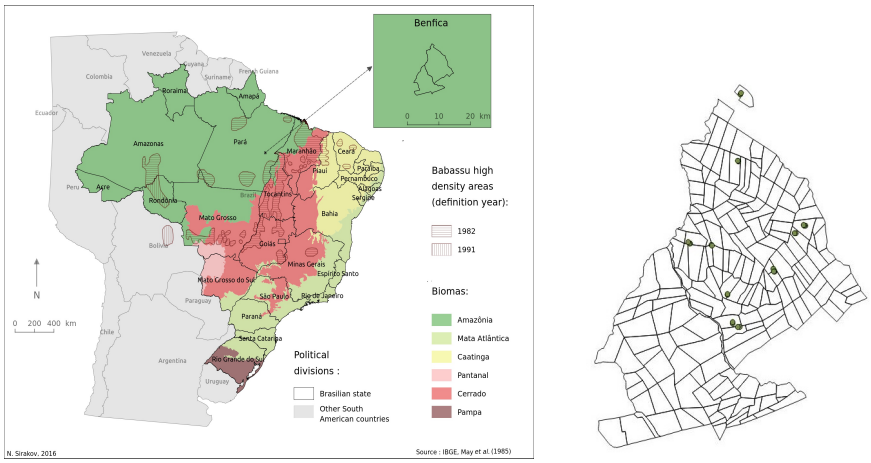
The fieldwork protocol was developed in early 2012. Seven surface areas of pastureland were designated for this research project. Within each piece

Fig. 1: Stage definition from Mitja et al. (2019)



of land, two to three transects (linear paths) were defined and visibly marked out on the ground dos Santos et al. (2017). We obtained the total number of seventeen transects from which we collected data regarding growth, sprouting and fruiting processes. In addition, we also used seven supplementary transects to obtain more data about the growth process, especially for stages 5 and 6

Fig. 2: PA Benfica in Brazil. Magnification of the localization of transects within the farmer's piece of land. The localization of some transects (green circle) are superimposed.



(Table 1).

Table 1: Design of the field sampling

Biological process	Stages	Nb of transects
Growth	1 to 4	17
	5 and 6	17 + 7
Sprouting	new stage 1	17
Fruiting	6	17

The counting and marking of all individuals at all six stages of development was performed within these transects. In 2013, 2,987 palm trees of all stages were identified. In 2014 and 2015, we surveyed the same palm trees, to which we respectively , added 471 and 482 newly recruited individuals of stage 1 (Table 2).

3. Model

The modelling was constructed based on three distinct experimental data sets (see flowchart Fig. 3) that were observed during two observation periods

Table 2: Empirical probabilities and sample size for babassu in pastures

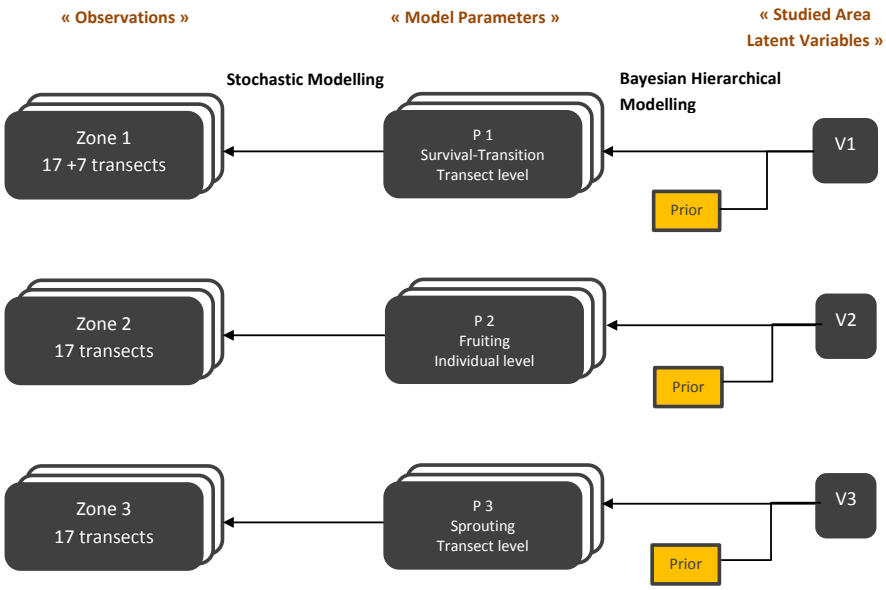
Stage		2013-2014	2014-2015
1	death	0.08	0.21
	idle	0.66	0.66
	transition	0.26	0.13
	Sample size	736	956
2	death	0.04	0.03
	idle	0.78	0.87
	transition	0.18	0.1
	Sample size	1737	1548
3	death	0.006	0
	idle	0.988	0.98
	transition	0.006	0.02
	Sample size	340	638
4	death	0	0
	idle	0.79	0.76
	transition	0.21	0.24
	Sample size	19	21
5	death	0	0
	idle	0.83	0.87
	transition	0.17	0.13
	Sample size	52	47
6	death	0.01	0.01
	idle	0.99	0.99
	Sample size	103	111

(one in 2013-2014 and one in 2014-2015) on 17 or 24 transects.

We decided to infer three independent models based on three different observational units for the following reasons:

- The transition-survival model was our main concern for answering demographic questions such as the existence of a bottleneck or a reverse J-shape curve. The survey was constructed for that purpose. The sample land was chosen to be representative of the different ages of pasture in Para. The 24 transects were chosen in order to sample a sufficient number of babassu

Fig. 3: Flowchart of the whole process



palm trees in different stages for the transition survival study, including 7 very specific transects focusing on old babassu palm trees (stage 5 and 6).

- The fruiting and sprouting processes were less precisely observed and only on the 17 transects. We wanted to avoid error propagation of the measurement or modelling approximations to the transition-survival modelling. Therefore, we decided to construct independent fruiting and sprouting models. Moreover, estimation of these models on the 17 transects might not be adapted to the 7 specific transects of older palm trees. However, even if the three models were independently estimated, they could be combined to simulate a life cycle.

- The survey did not focus on the seed dispersal (even if it exists, see Mitja and Ferraz (2001) for more detailed) or the germination. No information

was collected on the germination of fruits or seedlings. No information was available for outside the transects, (linear paths). Without this information, it was not possible to fit and use seed dispersal models such as the one proposed by Clark et al. (1999). Localization of adult palm trees around or adjacent to the transect would be necessary to achieve such modelling, and the new recruits were considered to be of unknown origin.

Brief description of the three independent models

- A demographic model with two levels (transect - studied area in Benfica) was defined. The observation were performed at an individual level, but the palm trees were considered to be independent, and the observations were pooled at the transect level for each stage. Demographic stochasticity was modelled with hierarchical Dirichlet (prior) and multinomial (likelihood) laws.

- The annual number of fruits from a palm tree was modelled with a log-Normal distribution to restrict them to positive values and to maintain flexibility in the meantime. A non-conjugated hierarchical model is used to model the annual mean number of fruits per palm tree across the studied area level (Normal prior) in addition to the observation/individual level (logNormal likelihood). This model is conditional on the probability that a palm tree can have fruits. This probability is modelled with a classical conjugate Beta-Binomial distribution.

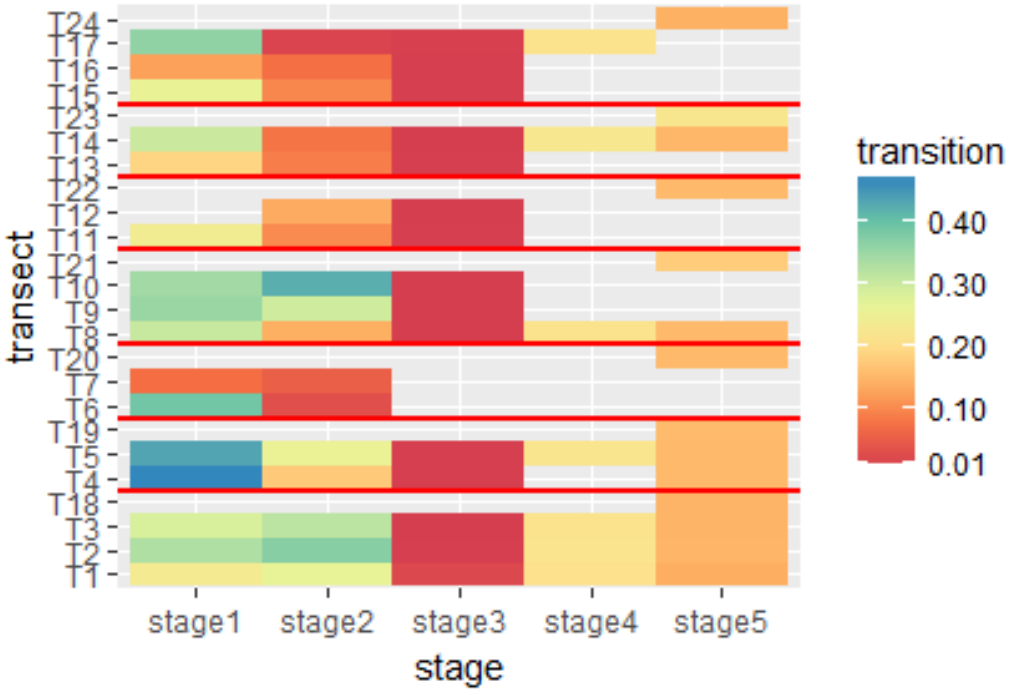
- We used a classical conjugate Beta-Binomial distribution to model the probability of sprouting across the studied area. Recall that the sprouting process was observed at the transect level, from a number of pooled observed fruits (including contribution from Babassu palm trees besides/along and outside the transect).

3.1. Hierarchy

The ecological process was observed on transects belonging to the designated area of land. Only two levels were considered for the transition survival process:

the transect and the whole studied area in Benfica. Moreover, our estimation concerning the probability of transition between stages (see Fig. 4) did not exhibit any structure at the level of the farmer's piece of land (each piece of land is separated by a red line). No spatial structure was observed, except for the transition from stage 2 to stage 3 where a link with spatial coordinates might be possible (see appendix and discussion) but it was difficult to estimate or ensure with our dataset. Therefore, we preferred to employ a multilevel model using the level of the transects and of the whole studied area in Benfica.

Fig. 4: Estimation of the probabilities of transition from one stage to the next per transect (transects are grouped by farmer's pieces of land, with each separated by an underlying red line)



3.2. Observation process and definition of the observed variables

Survival-transition process

Observations were performed on 17 (+7 for stages 5 and 6) transects (see Table 1). For transect l, we observed the total number of plants in stage i at year $t-1$, $n_{i,t-1,l}$. We also observed the number of survivors at year t in stage i, $n_{i,t,l}$. This number was the sum of the observed number of survivors that remained in this stage, denoted as $s_{i\to i,t,l}$ and of the observed number of survivors from stage $i-1$ who had moved to stage i, denoted as $s_{i-1\to i,t,l}$, $n_{i,t,l} = s_{i\to i,t,l} + s_{i-1\to i,t,l}$. Moreover, in a given transect l at time t, the number of deaths among those in stage i was the total number of plants in stage i from the previous year minus the number of survivors from the previous year that remained in stage i minus the survivors from the previous year that changed stages, $dead_{i,t,l} = n_{i,t-1,l} - s_{i\to i,t,l} - s_{i\to i+1,t,l}$. From these observations, we could infer proportions for each transect l. We assumed that the expected proportions remained constant over time and denoted $\pi_{i\to i,l}$ as the expected proportion of plants that remained at the same stage i, $\pi_{i\to i+1,l}$ as the expected proportion moving from stage i to the next stage $i+1$ and $1-\pi_{i\to i,l}-\pi_{i\to i+1,l}$ as the expected mortality proportion in stage i. These three parameters allowed us to define a categorical distribution for each transect l:

$$\pi_{i,l} = (\pi_{i\to i,l},\, \pi_{i\to i+1,l},\, 1-\pi_{i\to i,l}-\pi_{i\to i+1,l})$$

We also defined a categorical distribution at the level of the whole studied area:

$$\pi_i = (\pi_{i\to i}, \pi_{i\to i+1}, 1-\pi_{i\to i}-\pi_{i\to i+1})$$

No observations performed at this level, therefore $\pi_{i\to i}$ and $\pi_{i\to i+1}$ are latent variables.

Fruiting and sprouting processes.

Fruiting First, the fruit production per palm tree was observed on the 17 transects for each palm tree in stage 6. The reproduction process begins with the definition of the sexual state (no inflorescence, male or female inflorescence or both inflorescences). The babassu palm tree has a complex reproductive

process that is not well understood and has not been well measured. The female and male inflorescences change over undefined periods of time. Therefore the reproductive process has been simplified and resumed with the observed reproductive state variable $u_{j,t,l}$, which indicates the presence of fruits at year t on palm tree j from transect l. This variable was coded as 1 when fruits were present and 0 if they were not. We assumed that the expected proportion ϕ of palm trees of stage 6 in a reproductive state 1 was constant over time and was common to all transects.

When a palm tree j of year $t-1$ was in a reproductive state 1, its number of fruits, denoted $Fruits_{j,t-1,l}$, was the sum of the observed number of fruits on the palm tree and on the ground immediately under the palm tree. This observation was performed directly in the field but also validated later by using photography of the palm tree. We approximated fruit observations using a normal distribution with a mean parameter μ and a variance parameter σ^2.

Sprouting Next, the seedlings were not observed outside the transect, but only the number of recruited seedlings in stage one per transect l and per year t was measured in the field. We denoted this variable $r_{t,l}$, but the origin of these new palm trees was unknown. They could represent fruit from an adult palm tree (stage 6) adjacent to the transect. Therefore, the total number of fruits on transect l, year $t-1$ was counted (regardless of their origin), and we denoted this variable $TotalF_{t-1,l}$. We assumed that the recruitment occurred only in the first stage and that the expected proportion of fruit providing a new palm tree in one year was constant over time and was common to all transects. We denoted this parameter λ.

Remark: All variables were directly measured/observed on the ground by expert biologists. We assumed that there were no errors and our observation process was identity.

3.3. Lefkovitch Matrices for the ecological process

Six growth stages have been defined for babassu Mitja et al. (2019). Vector $\mathbf{n}_{t,l} = (n_{1,t,l}, \ldots, n_{6,t,l})'$ is the state vector, where $n_{i,t,l}$ is the number of individuals in stage i at year t in transect l. The ecological process is modelled with a Markovian process:

$$E\left(\begin{array}{c|c} n_{1,t,l} & n_{1,t-1,l} \\ \vdots & \vdots \\ n_{6,t,l} & n_{6,t-1,l} \end{array}\right) =$$

$$\begin{pmatrix} \pi_{1\to 1,l} & 0 & 0 & 0 & 0 & 0 \\ \pi_{1\to 2,l} & \pi_{22,l} & 0 & 0 & 0 & 0 \\ 0 & \pi_{2\to 3,l} & \pi_{3\to 3,l} & 0 & 0 & 0 \\ 0 & 0 & \pi_{3\to 4,l} & \pi_{4\to 4,l} & 0 & 0 \\ 0 & 0 & 0 & \pi_{4\to 5,l} & \pi_{5\to 5,l} & 0 \\ 0 & 0 & 0 & 0 & \pi_{5\to 6,l} & \pi_{6\to 6,l} \end{pmatrix} \begin{pmatrix} n_{1,t-1,l} \\ \vdots \\ n_{6,t-1,l} \end{pmatrix} + E\begin{pmatrix} r_{t,l} \\ \vdots \\ 0 \end{pmatrix} \quad (1)$$

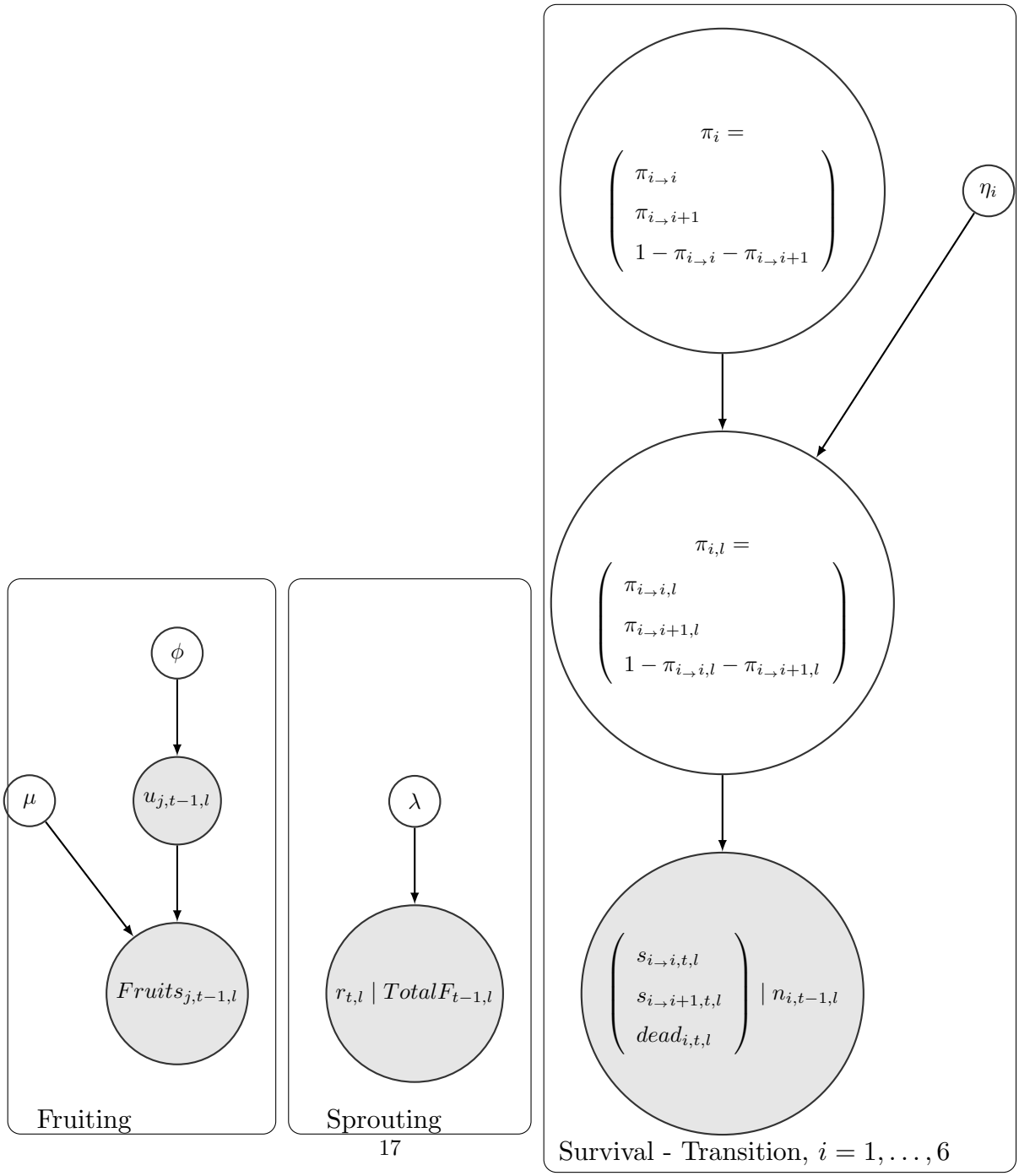
The projection of the population was modelled at the transect level with a Survival-Transition matrix following Caswell (2001).

The model must take into account the demographic stochasticity and hierarchy. Integration of different kinds/sources of variability was performed in a Bayesian framework. In this framework, all parameters except σ^2 were considered as random variables (see the DAG-Directed Acyclic Graph, Fig. 5, for details). This parameter was fixed in the observation in an empirical Bayes approach to ensure MCMC converged because we omitted the observation process error (lack of measurements allowing estimation of the observation error apart from the variability in the number of fruits) and also due to the very small sample size. Three independent Bayesian sub models were defined:

Survival-Transition for the six stages

According to Buckland et al. (2007), demographic stochasticity is modelled with multinomial laws (extension of the binomial law to a context with more than two possibilities).

Fig. 5: Representation of the hierarchy and dependences between variables with a Directed Acyclic Graph constructed for stage i and transect l. Three processes, fruiting, recruitment and survival-transition are represented from left to right. Observed variables are represented in grey.



- Evolution from year $t-1$ to year t of the six stages of transect l:

$$\begin{aligned}(s_{i\to i,t,l}, s_{i\to i+1,t,l}, dead_{i,t,l}) \mid \pi_{i,l}, n_{i,t-1} \sim Mult(n_{i,t-1,l}, \pi_{i,l}), i = 1,\ldots,5\\ (s_{6\to 6,t,l}, dead_{6,t,l}) \mid \pi_{6,l}, n_{6,t-1,l} \sim Mult(n_{6,t-1,l}, \pi_{6,l})\end{aligned} \tag{2}$$

To add to the hierarchy for each stage of the preceding model, we adapted a hierarchical Dirichlet distribution from Cowans (2004). Cowans proposed a hierarchical Dirichlet document model where each document (transect for us) is modelled using a multinomial distribution whose parameters are given a Dirichlet prior. The mean value of the Dirichlet prior is itself assumed to be random and given a Dirichlet hyperprior. This kind of modelling with Dirichlet priors and Dirichlet hyperpriors is used for topic models Wallach et al. (2009) or information retrieval from documents Haffari and Teh (2009).

- Hierarchical Dirichlet distribution to add to the hierarchy for each stage i

$$\begin{aligned}\pi_{i,l} \mid \pi_i, \eta_i &\sim Dirichlet(\eta_i \times \pi_i),\ i = 1,\ldots,6\\ \pi_i &\sim Dirichlet(Weights_i),\ i = 1,\ldots,6\\ \eta_i &\sim Uniform(u_1, u_2),\ i = 1,\ldots,6\end{aligned} \tag{3}$$

The shared mean parameter by stage i induces sharing of information across transects. $Weights_i$ of the Dirichlet hyperprior were calibrated with both observed proportions (as in an empirical Bayes approach) and uniform distribution (following Cowans (2004) idea) to test the sensitivity of the model to this hyperprior. The latent variable η_i could be considered to be a size effect which indicated the level of confidence in the realization of the latent vector π_i.

Fruiting and number of fruits.

The annual number of fruits of a palm tree in stage 6 is conditioned by its reproductive status $u_{j,t-1,l}$:

$$\begin{cases} Fruits_{j,t-1,l} \mid u_{j,t-1,l} = 1 \sim logNormal(\mu, \sigma^2)\\ Fruits_{j,t-1,l} \mid u_{j,t-1,l} = 0 \sim \delta_0\end{cases} \tag{4}$$

where δ_0 indicates the Dirac mass at 0. The logNormal distribution was preferred to the usual Poisson or negative binomial distributions to increase the flexibility of the modelling to address over dispersion and add explicitly to the population latent variable μ (the mean annual number of fruits inferred at the level of the studied area). The logNormal distribution provided the best fit our data (see supplementary materials, Fig. 10). Next, we assumed that the palm trees in stage 6 shared the same parameter of fruiting ϕ.

$$u_{j,t-1,l} \sim Bernoulli(\phi) \tag{5}$$

To add to the hierarchy allowing the sharing of information across palm trees (and transects), we defined:

$$\begin{aligned} \phi &\sim Beta(a, b) \\ \mu &\sim N(m, V) \end{aligned} \tag{6}$$

Sprouting (or recruitment in the first stage).
Recruitment in the first stage $r_{t,l}$ was stochastic and was defined Conditional on the total number of fruits present on a transect l, year $t-1$:

$$r_{t,l} \mid TotalF_{l,t-1} \sim Binomial(TotalF_{l,t-1}, \lambda) \tag{7}$$

To add to the hierarchy allowing the sharing of information across palm trees (and transects), we defined:

$$\lambda \sim Beta(c, d) \tag{8}$$

3.4. Hyperparameters

Two sets of hyperparameters were used, with one per year of estimation. For the first year, the prior was less informative as no expert knowledge was available. For the second year, we used the marginal posterior mean value and variance of the variables (ϕ, λ, π_i) to calibrate the hyperparameters. For both

years the hyperparameters m, V and σ^2 were calibrated from observed field data.

The values are presented in Table 3.

	Weights 1	2	3	4	5	6
Non info	(1/3,1/3,1/3)	...		...	(1/3,1/3,1/3)	(1/2,1/2)
Info	(.11,.63,.28)	(.04,.79,.17)	(.01,.98,.01)	(0,.79,.21)	(0,.83,.17)	(.01,.99)
	ϕ :(a,b)	λ :(c,d)	μ :(m, V)	σ^2	$\eta_i : (u_1, u_2)$	
Non Info	(1.5,1.5)	(1.5,1.5)	(313,4975)	59703	(1,200)	
Info	(13.5,12.5)	(472,3284)	(313,4975)	59703	(1,200)	

Table 3: Hyperparameter values for non-informative and informative priors (informative: hyperparameters were calibrated with the results of the first year posterior estimation)

3.5. Bayesian inference and model validation

In Bayesian frameworks, estimation of the variables of interest is performed for the posterior distribution. In our case, the variables of interest were for transition-survival $\theta_1 = ((\pi_{1,l}, \ldots, \pi_{6,l})_{l=1,\ldots,24}, \pi_1, \ldots, \pi_6, \eta_1, \ldots, \eta_6)$, recruitment in the first stage $\theta_2 = \lambda$ and fruiting $\theta_3 = (\phi, \mu)$. Whether the prior distribution of each θ_p is conditional/depends on the hyperparameter values. The vectors of those hyperparameters are $param_1 = (Weights_1, \ldots, Weights_6, u_1, u_2, \mathbf{n}_{t-1})$, $param_2 = (c, d, TotalF_{t-1,1}, \ldots, TotalF_{t-1,17})$ and $param_3 = (a, b, m, V, \sigma^2)$. Each vector is considered fixed, and its values are given in Table 3. Summaries for $(\mathbf{n}_{t-1,l})_{l=1,\ldots,24}$ and the histogram for $(TotalF_{t-1,l})_{l=1,\ldots,17}$, are given in Table 2 and Supplementary Material, Fig. 10, respectively. Observations or data respectively are the realization of the following variables:
$data_1 = (s_{1\to 1,t,l}, s_{1\to 2,t,l}, \ldots, s_{6\to 6,t,l})_{l=1,\ldots,24}$, $data_2 = (r_{t,l})_{l=1,\ldots,17}$ and $data_3 = (u_{j,t-1,l}, Fruits_{t-1,l})_{j\in transect_l, l=1,\ldots,17}$.
Based on those definitions, from Bayes' rule, the posterior distribution is proportional to

$$p(\theta_p \mid data_p, param_p) \propto p(data_p \mid \theta_p, param_p)\, p(\theta_p \mid param_p) \qquad (9)$$

To simplify the notation, we will omit the dependence on *param* in the following equations. Of course, we are clearly aware that posterior estimation is conditional on *param* and therefore sensitive to the values of the hyperparameters, and this point will be discussed and studied later.

Using the DAG (Fig. 5) and conditional dependencies, Equation (9) can be decomposed into a product, depending on the type of observation: Transition-Survival, Fruiting and Sprouting:

$$
\begin{aligned}
&p(\theta_1 \mid data_1) \propto \\
&p(\text{Survival data} \mid (\pi_{1,l}, \ldots, \pi_{6,l})_{l=1,\ldots,24},\ \pi_1, \ldots, \pi_6, \eta_1, \ldots, \eta_6) \times \\
&p((\pi_{1,l}, \ldots, \pi_{6,l})_{l=1,\ldots,24},\ \pi_1, \ldots, \pi_6, n_1, \ldots, n_6) = \\
&\Pi_i \Pi_l\ p(s_{ii,t,l}, s_{i\,i+1,t,l}, dead_{i,t,l} \mid \pi_{i,l})\, p(\pi_{i,l} \mid \pi_i, \eta_i)\, p(\pi_i)\, p(\eta_i)
\end{aligned}
$$

$$
\begin{aligned}
&p(\theta_2 \mid data_2) \propto \\
&p(\text{Sprouting data} \mid \lambda) \times p(\lambda) = \\
&\Pi_l\ p(r_{t,l} \mid \lambda)\, p(\lambda)
\end{aligned}
$$

$$
\begin{aligned}
&p(\theta_3 \mid data_3) \propto \\
&p(\text{Fruiting data} \mid \mu, \phi) \times p(\mu, \phi) = \\
&\Pi_l \Pi_{j \in transect_l}\ p(Fruits_{j,t-1,l} \mid \mu, u_{j,t-1,l})\, p(u_{j,t-1} \mid \phi)\, p(\phi)\, p(\mu)
\end{aligned}
$$

Sensitivity to Informative versus Non-Informative priors in a learning approach

Estimation based on the first year of observation was performed according to Helser et al. (2012) since no study to date has provided prior information on babassu growth or reproduction processes in the babassu environment. We endorsed the approach of Smith and Wakefield (1994), who suggest that priors should be non-informative to ensure that the likelihood dominates the prior. The hyperparameters characterizing the distribution of $\pi_i, \eta_i, \phi, \lambda$ and μ were chosen to provide a prior that was as flat and non-informative as possible. Then, in a learning approach, the value of the hyperparameters $Weights_i$ was set to the posterior estimates of π_i. We compared the estimation based on the second

year of observation obtained with each set of hyperparameters. This comparison allowed us to study the sensitivity of the model and its stability over the years.

Model Validation

The predictive posterior distribution can be used to check whether the model is consistent with the observed data Gelman et al. (1996) measurement, y^{obs} is the observed value and y^{rep} is a replicate sampled from the predictive posterior $\int p(y^{rep} \mid \theta)\, p(\theta \mid data)\, d\theta$ Steinbakk and Storvik (2009). Posterior p-values of approximately 0.5 indicate that the distributions of replicated and actual data are close, while values near zero or one indicate a discrepancy between them.

Moreover, validation was possible by leaving out 25% of the transects for the second year of observation 2014-2015. The posterior population estimate π_i was then used to predict the total expected number of survivors $(s_{i\rightarrow i}, s_{i\rightarrow i+1}, dead_i)$ of the omitted transects. The total number per stage of the omitted transects was considered to be consistent with a prediction at the level of the studied area.

Computation with R and JAGS

JAGS which uses the dialect of BUGS Lunn et al. (2000), Plummer (2003), Lunn et al. (2013) was used to fit the model to the data with the package rjags. The estimates of parameters were evaluated based on 200,000 samples after 50,000 burn-in samples, thinning to one draw for each sample, by MCMC simulation of the joint posterior distribution. The MCMC inference samples analysis was processed with the coda package Plummer et al. (2006). Convergence diagnostics consisted of visuals (traces and autocorrelations) and statistical calculations using the test of Gelman et al. (2003) based on Gelman and Rubin (1999). For all variables in the model, the Gelman criteria of between-chain variance over within-chain variance converged into one. More information on the convergence diagnosis and on the script of the model are available in the supplementary materials.

4. Results and Discussion

4.1. Model validation

We focused the validation analysis on the Survival-Transition matrix at the level of the studied area. The validation concerned the second year of observation 2014-2015 with the informative prior.

First, we computed the posterior predictive p-value for estimation of the survival-transition probabilities obtained for the second year of observation (2014-2015) and with the informative prior (see Table 4). When the ppp-value was much smaller than 0.5 (e.g., 0 or 0.1), then the probability simulated with the model was less variable than the observed probability, and the model failed to reproduce the variability of the data. Thus, the model underestimated the data variability. We imputed this phenomenon as a higher variability observed in the data during 2014-2015 than during 2013-2014, which may indicate a year effect. A longer time series of observation may well be needed to define an informative prior, as suggested by Bierzychudek (1999).

	ppp-values		
Stage i	$1-\pi_{i\to i}-\pi_{i\to i+1}$	$\pi_{i\to i}$	$\pi_{i\to i+1}$
1	0	0.5	0.22
2	0.62	0.1	0.04
3	0.39	0.14	0.24
4	0.15	0.64	0.76
5	0.14	0.43	0.57
6	0.71	0.36	-

Table 4: Posterior predictive p-values of the survival-transition probabilities (π_i, $i = 1, \ldots, 6$) obtained for the second year of observation 2014-2015 and with the informative prior

Next, we checked the adequacy of the model with the cross-validation approach (see Table 5) for the second year of observation 2014-2015 and with the informative prior. The young stages (namely one and two) showed a poor

prediction, as the observed values were not in the 95% credible interval of the prediction for all variables. Again, this mismatch for the young stages was a result of the extreme variability observed in the field for the dynamics of the young during the two periods of observation (2013-2014 and 2014-2015).

	$dead_{i,2015}$		$s_{i\to i,2015}$		$s_{i\to i+1,2015}$	
Stage i	obs.	pred.	obs.	pred.	obs.	pred.
1	46	$27_{[25,40]}$	75	$81_{[65,96]}$	19	$32_{[20,46]}$
2	33	$31_{[16,51]}$	481	$443_{[410,471]}$	39	$79_{[74,107]}$
3	0	$0_{[0,2]}$	83	$82_{[79,83]}$	0	$1_{[0,3]}$
4	0	$0_{[0,1]}$	3	$4_{[2,5]}$	2	$1_{[0,3]}$
5	0	$0_{[0,1]}$	8	$9_{[6,10]}$	2	$1_{[0,4]}$
6	1	$1_{[0,3]}$	58	$58_{[56,58]}$	-	-

Table 5: Cross-validation for the year 2015 and for each stage $i = 1, \ldots, 6$, the observed (*obs.*) and predicted (*pred.*) numbers of palm trees that died ($dead_{i,2015}$), survived in the same stage ($s_{i\to i,2015}$) and survived and evolved toward the upper stage ($s_{i\to i+1,2015}$). The credible interval at a 95% level was computed for each prediction.

Eventually, the learning approach of using the posterior of the first period 2013-2014 to calibrate the prior of the second period 2014-2015 was not satisfactory for the young stage (namely 1 and 2).

4.2. Survival and transition between states: identification of a bottleneck

Table 6 summarizes the estimates of the survival-transition probabilities for both periods (2013-2014 and 2014-2015). The dynamic can be divided into two parts. Before compared with after stage 3, mortality was significantly elevated (5% to 27%). The first stage exhibited a high year variability, as underlined in the model validation section. This competition disappeared at a later stage of palm tree growth. The death probability for the first stage was estimated at 12% for the first period and 27% for the second period. These values were both lower than the estimate of 38% published by Anderson (1983) and Anderson and Balick (1988) concerning another experiment in Brazil. After stage 3,

the estimates were not sensitive to the prior or the year effect. The mortality rates were lower and difficult to estimate with small sample sizes. Except for the transition between stage 3 and stage 4 (lower than 1%), the probability of transitions were significant (10% to 28%). Hence, these results showed a clear bottleneck at stage 3. Due to its extremely low probability, the transition between stage 3 and 4 had a large impact on the "needed" time for a babassu to reach the adult stage 6. Thus, a supplementary sampling effort should be performed for stages 3 and 4 to better estimate and understand this transition. Stage 3 is critical in the biology of the babassu because the palm terminal bud rises from the ground during this stage.

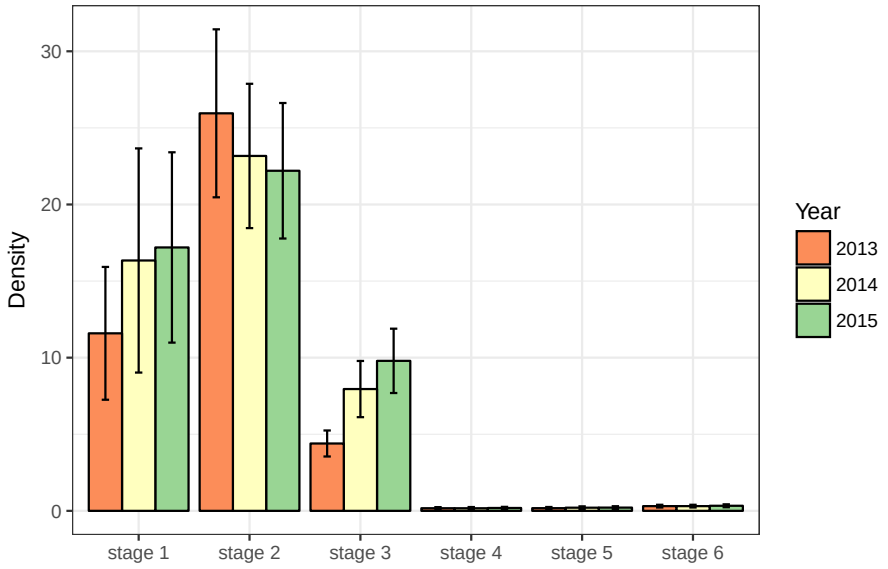
Table 6: Sensitivity analysis of the prior. The given estimates are posterior mean values for the transition-survival probabilities.

Probabilities		2014 non info.	2015 non info.	2015 info.
$\pi_{1,t}$	Death	0.12	0.27	0.21
	Idle	0.6	0.52	0.57
	Transition	0.28	0.21	0.22
$\pi_{2,t}$	Death	0.05	0.07	0.06
	Idle	0.79	0.83	0.81
	Transition	0.16	0.10	0.13
$\pi_{3,t}$	Death	0.005	0.001	0.005
	Idle	0.993	0.994	0.989
	Transition	0.002	0.005	0.006
$\pi_{4,t}$	Death	0.02	0.02	0.01
	Idle	0.77	0.75	0.77
	Transition	0.21	0.23	0.22
$\pi_{5,t}$	Death	0.01	0.01	0
	Idle	0.85	0.87	0.85
	Transition	0.14	0.12	0.15
$\pi_{6,t}$	Death	0.01	0.01	0.01
	Idle	0.99	0.99	0.99

4.3. Link between the bottleneck and population structure

Finally, we observed a bottleneck at the end of stage 3 in pastures. This bottleneck was observed in a primary forest by Barot et al. (2005). Moreover, a comparison of our bar plot of densities by stage (see Fig. 6) revealed a shape similar to Barot et al. (2005) for stages 1 to 3 (corresponding to ELS, SLS and Juv1, respectively, in their article); however in the next stage situated after the bottleneck, the densities are collapsed toward zero in our case.

Fig. 6: Mean and standard error of the transect densities by stage and by year of observation.



We expected, a priori, high densities in all young stages for this palm specie,s which is considered to be an invasive plant species in pastures, Barot et al. (2005). A pasture is an open field/environment with less competition for light or with other species. In open areas, the number of germinations per fruit is lower but growth in stage 1 is more rapid, Mitja and Ferraz (2001).

Fig. 6 shows an increase in density in stage 3 over the time. In addition, the mortality rates for stages 3 to 6 were very low during this period. Therefore, we may assume that the observed bottleneck was due to the low transition rate

from stage 3 to stage 4. Multiple causes could have contributed or explained this low transition.

First, competition for light and nutrients with fodder plants or introduced grasses might not be negligible in the younger stages. As stated in the article by Sankaran et al. (2004), "tree grass coexistence in savannas", competition-based models or demographic-bottleneck models, or a mixture of the two, could explain the population structure of babassu in the pasture.

Second, the management of the pasture, with human intervention, has an impact on what is believed to be an appropriate density of babassu palm trees for cattle shelter. Runk (1998) observed this kind of human intervention in the regulation of Ivory palm tree, where "the lack of juveniles to the sub adult stage appears to be a population bottleneck in managed stands, limiting the number of adults". However, during the period of our study, management was limited to an annual clearing to respect our research on the population dynamics and estimation of the natural mortality. Annual clearing of pasture plots by cutting leaves from stages 2 to 4 and sometimes even stage 5 can delay the transition between stages because the palm tree must reinvest in the production of new leaves, Mitja et al. (2018).

Third, the impact of cattle (for example, trampling, grazing) might provide an explanation for a demographic bottleneck model. The terminal meristem of the plant is below ground until stage 3, and the plant can regenerate even if the aerial part is damaged (grazed, etc.).These non-lethal actions (as we estimated a very low mortality in stage 3) might delay the transition and thus create a bottleneck for the transition from stage 3 to stage 4.

Fourth, none of these suggestions explain why this bottleneck or at least a reverse J-shaped curve (with a lower density in stage 4) was observed in primary rain forests. In addition to the competition for light with other trees (or babassus), the population structure could be due to different times of residence in the different stages. The prolonged juvenile phase of trees is obviously a central component of their life strategy. Brunner and Ove (2004) have suggested a different mechanism (mainly related to chromatin-based repression) to explain

this prolongation. One hypothesis is that epigenetic mechanisms play a central role in the regulation of the tree phases. Bräutigam et al. (2013) have explored the epigenetic variation and adaptive capacity of plant species, including forest trees. Considered in this way, the prolonged juvenile phase might be a contributory strategy to the babassu's ability to colonize or persist in pastures.

4.4. Other biological information: fruiting and sprouting

The fruiting and sprouting parts of the model were simplifications of the complex reality. Nevertheless, rate estimation of fruit production (55%) and sprouting (12%) exhibited good statistical properties, and these estimations were enriched by the Bayesian learning period. We suggest that these processes (fruiting and sprouting) remained unchanged during the two periods of observation.

Fruiting

The mean estimate values $\pm$ the standard error of the fruiting rate (ϕ) were 0.52 ± 0.10 for 2013 and 0.54 ± 0.07 for 2014. The estimation for 2014 was performed with an informative prior using a Bayesian learning approach (the posterior of the first period was used as the prior for the second period). Biologically speaking, these values indicate that a little more than fifty percent of the adult palm trees contribute annually to the fruiting process.

The mean estimate values $\pm$ the standard error of the number of fruits per palm tree (μ) respectively were 313 ± 50 for 2013 and 317 ± 48 for 2014.

Sprouting

The mean value estimates for the sprouting probability λ were approximately ≈ 0.12 for both periods, with standard error of 0.005 for the first period 2013-2014 and 0.004 for the second period 2014-2015. λ represents the probability that a fruit produced in year t becomes a seedling (stage 1) in the next year $t+1$.

4.5. Bayesian Learning

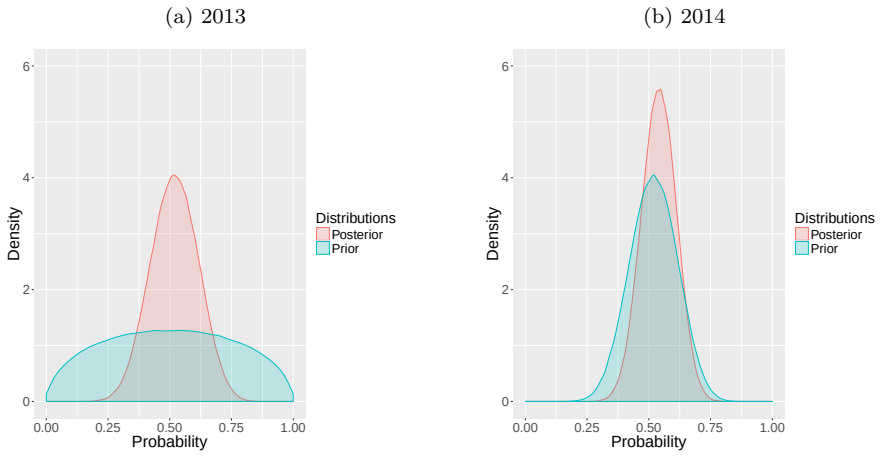
Fruiting/sprouting

The estimation for the first period 2013-2014 was performed with a non-informative prior. A flat distribution was used as illustrated in blue, for the fruiting rate ($\phi \sim Beta(a = b = 1.5)$), Fig. 7a. The distribution obtained *a posteriori* (in red) was concentrated on the estimated mean value of 54%. This figure illustrated the Bayesian learning approach with the evolution between the flat distribution (non-informative prior) to the conditional distribution (knowledge of the information provided by the data). This informative distribution (in red, Fig. 7a) obtained during the first period of observation was then used as the prior (in blue, Fig. 7b) for estimation of the fruiting rate during the second period of observation, 2014-2015. Again, the distribution obtained *a posteriori* (in red) exhibited a concentration around the same mean estimated value. Bayesian learning helped to increase the precision of the estimation for the fruiting rate. We observed similar figures for the sprouting rate and the number of fruits per palm tree. Concerning transition-survival probabilities, this learning approach was not so fruitful, because of important variability-fluctuation between the two periods of observation mainly during the early stages (1 and 2).

5. Conclusion and perspectives

To better understand the babassu population dynamics in pastures, we used a Bayesian hierarchical matrix model. This model with two levels of hierarchy (studied area in Benfica and transect) allowed a more reliable estimation of the population dynamics because information is shared between transects to improve the estimation at the level of the studied area and then retrospectively at the transect level. This is a significant improvement over direct empirical estimations using small sample sizes, with missing data and an unbalanced design, Fontez and Cavalli (2014). In our case, the observed stages in transects were unbalanced, as shown in Fig. 4. The model did not reveal unobserved empirical facts but provided a direct and reliable estimation at the level of the

Fig. 7: Estimation of the fruiting rate. Prior and posterior for both periods.



studied area. This model enabled us to demonstrate a bottleneck at stage 3 in the population dynamics. The origin of this bottleneck is not known but was not due to mortality (low values). The bottleneck between stages 3 and 4 has significant consequences for the population dynamics of babassu, but in the present state of knowledge provided by the model, it is difficult to specify the exact consequences of this bottleneck. In fact, two extreme cases can occur: 1. all the palms have an equal probability of passing from stage 3 to stage 4 each year; 2. only a portion of the palm trees are able to undergo this transition.

- In the first case, all the palm trees are impacted by the bottleneck and a rough estimation of the period in stage 3 with a geometrical model could be $\pi_{3\to 4}/(1-\pi_{3\to 3})^2 = 0.005/(1-0.99)^2 \approx 50$ years. This bottleneck dramatically raised the age at which a palm tree becomes an adult. This result contradicts information from experiments (some palm trees became adults only after 10-12 years).

- In the second case, a fraction of individuals f can successfully transit and a fraction $1-f$ will never transit. Therefore, $\pi_{3\to 4}$ is a compromise (a mean value) between those two kinds of palm trees. The very low

value estimated for $\pi_{3\to 4}$ could be due to this fraction $1-f$ that never transited. According to Barot et al. (2005) fraction f that transits is probably composed of "juvenile with a high survival and growth rate to be quickly recruited into the adult stage and hence representative of a small percentage of the population". For this kind of juvenile, using a geometrical model, we deduced the expected time spent in the stages that are above soil and observable (4-6). Using estimates presented in Table 6 we obtained (3.68 to 4.16) for stage 4 and (6.22 to 7.1) for stage 5 for a total sum of these stages of 9.9 to 11.3 years. In the field, the people interviewed suggested 10 to 12 years. If we consider the stages that are above the soil and observable (4-6), the predictions are coherent with our model estimations.

The reality is undoubtedly between these two extreme cases. To arbitrate between these two hypotheses, follow-up sequences should be available in the future, which will soon be possible as new surveys are underway.

Concerning the second case, one possible explanation could be the high genetic variability of the babassu, which is a wild plant. Thus, important genetic-environment interactions may be involved. A genetic difference for the f fraction of successful babassu is coherent with the observed increase in density in stage 3 over the years. At the population level, clearing of the forest and the use of fire to clear brush have resulted in a population bottleneck. Following Rapp and Wendel (2005), these anthropic actions may have created genomic stress leading to the induction of epigenetic phenomena. From an epigenetic perspective, a population bottleneck leads to a decrease in genetic diversity but an increase in phenotypic variability. In fact, we observed a great density variability between pieces of land, which could be due to phenotypic variability (see, for example, the probability of transition from stage 2 to 3 in the supplementary material). Based on our study of very local and limited points of observation in Pará, it is impossible to comment further without additional research.

Another explanation could be the past history of each plant. New sequences

would make it possible to determine for an individual whether the past history of the passages from stage 1 to stage 2 and stage 2 to stage 3 affected the probability of passage from stage 3 to stage 4.

In open environments (such as pastureland), the number of seedlings per fruit is lower for babassu, Mitja and Ferraz (2001). In the present work, we highlighted a high and year-dependant mortality rate at stage 1 and a bottleneck at stage 3. These three points highlight a new perspective maybe in contradiction with the common idea that the babassu is an invasive species, Barot et al. (2005). These three points in conjunction with harvesting might explain the negative evolution of density in pastures, observed in Mitja et al. (2018), and the need for reliable preventive scientific information to help manage babassu population dynamics in pastures.

Our model could be extended to a larger scale (region, country) by incorporating palm density information from high spatial resolution satellite images dos Santos et al. (2017). However, to achieve the sustainable management of the species, a more comprehensive study of the mechanisms underlying the low transition rate between stage 3 to 4 is needed. Some factors, such as competition for light and nutrients with fodder plants as well as information on the genetic diversity (or a potential epigenetic mechanism) should first be adressed.

Author's contributions

Nikolay Sirakov drafted this article, and his PhD thesis is on this work. B. Fontez and P. Loisel provided the modelling and co-drafted the article. A. dos Santos, provided the field data and proof-read the article. D. Mitja drafted the article and helped to define the babassu stages and interpret the results. T. Libourel contributed to the modelling and implementation.

Acknowledgments

This PhD research work has been accomplished as a part of the Open Science research project (under the reference ID 1403-026) financed by the Agropo-

lis Foundation and by the French National Institute of Agronomical Research (INRA) and Montpellier SupAgro. We thank all our partners who participated in the Open Science research project (especially É. Delaître and L. Demagistri from IRD and our Brasilian partners, with special thanks to I. de Souza Miranda and D. Deurival da Costa Carvalho) and E. Bonnier for providing language support, proof-reading and reformulation.

References

Anderson, A.B., 1983. The biology of Orbignya Martiana (Palmae), a tropical dry forest dominant in Brazil. Ph. D.. University of Florida.

Anderson, A.B., Balick, M.J., 1988. Taxonomy of the babassu complex (Orbignya spp. palmae). Systematic Botany 13, 32–50.

Anderson, A.B., May, P.H., Balick, M.J., 1991. The Subsidy from Nature: Palm Forests, Peasantry, and Development on an Amazon Frontier. Columbia University Press, New York.

Anthelme, F., Lincango, J., Gully, C., Duarte, N., Montúfar, R., 2011. How anthropogenic disturbances affect the resilience of a keystone palm tree in the threatened andean cloud forest? Biological Conservation 144, 1059 – 1067. The New Conservation Debate: Beyond Parks vs. People.

Araújo, F.R., Lopes, M.A., 2012. Diversity of use and local knowledge of palms (Arecaceae) in eastern Amazonia. Biodiversity and Conservation 21, 487–501.

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R.M., Thomson, J.R., Ferraz, S.F.d.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro de Castro Solar, R., Vieira, I.C.G., Arago, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., Jr, R.C.d.O., Souza Jr, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-de Mello, F.Z., Veiga, R.C.S., Venturieri, A., Gardner, T.A., 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature 535, 144.

Barot, S., Gignoux, J., Vuattoux, R., Legendre, S., 2000. Demography of a Savanna Palm Tree in Ivory Coast (Lamto): Population Persistence and Life-History. Journal of Tropical Ecology 16, 637–655.

Barot, S., Mitja, D., Miranda, I., Meija, G., Grimaldi, M., 2005. Reproductive plasticity in an Amazonian palm. Evolutionary Ecology Research 7, 1051–1065.

Bierzychudek, P., 1999. Looking Backwards: Assessing the Projections of a Transition Matrix Model. Ecological Applications 9, 1278–1287.

Bräutigam, K., Vining, K.J., Lafon-Placette, C., Fossdal, C.G., Mirouze, M., Marcos, J.G., Fluch, S., Fraga, M.F., Guevara, M.A., Abarca, D., Johnsen, O., Maury, S., Strauss, S.H., Campbell, M.M., Rohde, A., Díaz-Sala, C., Cervera, M.T., 2013. Epigenetic regulation of adaptive responses of forest tree species to the environment. Ecology and evolution 3, 399–415.

Brunner, A.M., Ove, N., 2004. Revisiting tree maturation and floral initiation in the poplar functional genomics era. New Phytologist 164, 43–51.

Buckland, S.T., Newman, K.B., Fernndez, C., Thomas, L., Harwood, J., 2007. Embedding Population Dynamics Models in Inference. Statistical Science 22, 44–58.

Buckley, Y.M., Briese, D.T., Rees, M., 2003. Demography and management of the invasive plant species Hypericum perforatum. II. Construction and use of an individual-based model to predict population dynamics and the effects of management strategies. Journal of Applied Ecology 40, 494–507.

Bullock, S.H., 1980. Demography of an Undergrowth Palm in Littoral Cameroon. Biotropica 12, 247–255.

Campos, J., Albuquerque, U., Peroni, N., de Lima Araújo, E., 2017. Population structure and fruit availability of the babassu palm (attalea speciosa mart. ex spreng) in human-dominated landscapes of the northeast region of brazil. Acta Botanica Brasilica 31, 267–275.

Caswell, H., 1989. Analysis of life table response experiments I. Decomposition of effects on population growth rate. Ecological Modelling 46, 221–237.

Caswell, H., 2001. Matrix population models : construction, analysis, and interpretation. Sinauer, Sunderland (Massachusetts).

Clark, J.S., 2003. Uncertainty and Variability in Demography and Population Growth: A Hierarchical Approach. Ecology 84, 1370–1381.

Clark, J.S., Silman, M., Kern, R., Macklin, E., HilleRisLambers, J., 1999. Seed dispersal near and far: Patterns across temperate and tropical forests. Ecology 80, 1475–1494.

Condit, R., Sukumar, R., Hubbell, S., Foster, R., 1998. Predicting population trends from size distributions: A direct test in a tropical tree community. The American Naturalist 152, 495–509.

Cowans, P.J., 2004. Information Retrieval Using Hierarchical Dirichlet Processes, in: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, New York, NY, USA. pp. 564–565.

Eacker, D., Lukacs, P., M. Proffitt, K., Hebblewhite, M., 2017. Assessing the importance of demographic parameters for population dynamics using Bayesian integrated population modeling. Ecological Applications 27, 1280–1293.

Enright, N., Watson, A.D., 1992. Population dynamics of the nikau palm, Rhopalostylis sapida (Wendl. et Drude), in a temperate forest remnant near Auckland, New Zealand. New Zealand Journal of Botany 30, 29–43.

Fearnside, P.M., 2005. Deforestation in brazilian amazonia: History, rates, and consequences. Conservation Biology 19, 680–688.

Fontez, B., Cavalli, L., 2014. Bayesian hierarchical model used to analyze regression between fish body size and scale size: application to rare fish species zingel asper. Knowl. Managt. Aquatic Ecosyst. , 10.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, Second Edition. 2 ed., Chapman and Hall/CRC, Boca Raton, Fla.

Gelman, A., Meng, X.L., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. Statistica sinica 6, 733–760.

Gelman, A., Rubin, D.B., 1999. Evaluating and using statistical methods in the social sciences - A discussion of "A critique of the Bayesian information criterion for model selection". Sociological Methods & Research 27, 403–410.

Guarín, J.R., del Valle, J.I., 2014. Modeling the stipe growth of the Oenocarpus bataua palm in the Central Cordillera of the Andes, Colombia. Forest Ecology and Management 314, 141–149.

Haffari, G., Teh, Y.W., 2009. Hierarchical Dirichlet Trees for Information Retrieval, in: Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, Boulder, Colorado, USA. pp. 173–181.

Hall, P., Bawa, K., 1993. Methods to assess the impact of extraction of nontimber tropical forest products on plant populations. Economic Botany 47, 234–247.

Helser, T.E., Lai, H.l., Black, B.A., 2012. Bayesian hierarchical modeling of Pacific geoduck growth increment data and climate indices. Ecological Modelling 247, 210–220.

Hobbs, T.N., Hooten, M., 2015. Bayesian Models - A Statistical Primer for Ecologists. Princeton University Press, Princeton.

Holm, J.A., Miller, C.J., Cropper, W.P., 2008. Population Dynamics of the Dioecious Amazonian Palm Mauritia flexuosa: Simulation Analysis of Sustainable Harvesting. Biotropica 40, 550–558.

Jha, A., 2010. World's forests can adapt to climate change, study says. The Guardian URL: https://www.theguardian.com/environment/2010/nov/11/climate-change-forests-water-amazon.

Johnson, E., Miyanishi, K., Kleb, H., 1994. The hazards of interpretation of static age structures as shown by stand reconstructions in a pinus contorta – picea engelmannii forest. Journal of Ecology 82, 923–931.

Lomnicki, A., 1999. Individual-based models and the individual-based approach to population ecology. Ecological Modelling 115, 191–198.

Lunn, D., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D., 2013. The BUGS book: A practical introduction to Bayesian analysis. CRC press.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10, 325–337.

Mitja, D., Delaître, E., dos Santos, A.M., de Souza Miranda, I., Coelho, R., Macedo, D., Demagistri, L., Petit, M., 2018. Satellite images combined with field data reveal negative changes in the distribution of babau palm after clearing off amazonian forest. Journal of Environmental Management 61, 321–336.

Mitja, D., Ferraz, I.D.K., 2001. Establishement of Babassu in Pastures in Pará, Brazil. PALMS 45, 138–147.

Mitja, D., Sirakov, N., dos Santos, A.M., González-Pérez, S.E., Macedo, J.D., Delaître, E., Demagistri, L., Loisel, P., Miranda, I.d.S., Rey-Valette, H., da Rocha, M.R.T., Fontez, B., Libourel, T., 2019. Viability of the babassu palm eco-sociosystem in Brazil: the challenges of coviability, in: Coviability of Social and Ecological Systems : Reconnecting Mankind to the Biosphere in an Era of Global Change. Springer-Verlag, p. accepted.

Norden, N., Chazdon, R.L., Chao, A., Jiang, Y.H., Vílchez-Alvarado, B., 2009. Resilience of tropical rain forests: tree community reassembly in secondary forests. Ecology Letters 12, 385–394.

Olmsted, I., Alvarez-Buylla, E.R., 1995. Sustainable Harvesting of Tropical Trees: Demography and Matrix Models of Two Palm Species in Mexico. Ecological Applications 5, 484–500.

Parent, E., Rivot, E., 2012. Introduction to Hierarchical Bayesian Modeling for Ecological Data. Chapman and Hall/CRC.

Pinard, M., 1993. Impacts of Stem Harvesting on Populations of Iriartea deltoidea (Palmae) in an Extractive Reserve in Acre, Brazil. Biotropica 25, 2–14.

Pinero, D., Martinez-Ramos, M., Sarukhan, J., 1984. A Population Model of Astrocaryum Mexicanum and a Sensitivity Analysis of its Finite Rate of Increase. Journal of Ecology 72, 977–991.

Plummer, M., 2003. JAGS: A program for analysis of bayesian graphical models using gibbs sampling, in: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20-22, Vienna, Austria.

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6, 7–11.

Rapp, R.A., Wendel, J.F., 2005. Epigenetics and plant evolution. New Phytologist 168, 81–91.

Richards, C.L., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., Durka, W., Engelhardt, J., Gaspar, B., Gogol-Döring, A., Grosse, I., van Gurp, T.P., Heer, K., Kronholm, I., Lampei, C., Latzel, V., Mirouze, M., Opgenoorth, L., Paun, O., Prohaska, S.J., Rensing, S.A., Stadler, P.F., Trucchi, E., Ullrich, K., Verhoeven, K.J.F., 2017. Ecological plant epigenetics:

Evidence from model and non-model species, and the way forward. Ecology Letters 20, 1576–1590.

Runk, J.V., 1998. Productivity and sustainability of a vegetable ivory palm (phytelephas aequatorialis, arecaceae) under three management regimes in northwestern ecuador. Economic Botany 52, 168–182.

Sampaio, S.M.N., 2008. Dinâmica e complexidade da paisagem do Projeto de Assentamento Benfica, Sudeste Paraense. Ph. D.. Universidade Rural da Amaznia. Belém, PA.

Sankaran, M., Ratnam, J., Hanan, N.P., 2004. Tree-grass coexistence in savannas revisited-insights from an examination of assumptions and mechanisms invoked in existing models. Ecology Letters 7, 480–490.

dos Santos, A.M., Mitja, D., Delaître, E., Demagistri, L., de Souza Miranda, I., Libourel, T., Petit, M., 2017. Estimating babassu palm density using automatic palm tree detection with very high spatial resolution satellite images. Journal of Environmental Management 193, 40–51.

dos Santos, A.M.d., Mitja, D., 2011. Wooded Cattle Pasture in the Benfica Seetling Project in Itupiranga, Para, Brazil. Revista Arvore 35, 919–930.

Schaub, M., Abadi, F., 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152, 227–237.

Smith, A., Wakefield, J., 1994. The hierarchical Bayesian approach to population pharmacokinetic modelling. International Journal of Bio-Medical Computing 36, 35–42.

Steinbakk, G.H., Storvik, G.O., 2009. Posterior Predictive p-values in Bayesian Hierarchical Models. Scandinavian Journal of Statistics 36, 320–336.

Tuljapurkar, S., Caswell, H. (Eds.), 1997. Structured-Population Models in Marine, Terrestrial, and Freshwater Systems. Springer US, Boston, MA.

Uchmaski, J., Grimm, V., 1996. Individual-based modelling in ecology: what makes the difference? Trends in Ecology & Evolution 11, 437–441.

Wallach, H.M., Mimno, D.M., McCallum, A., 2009. Rethinking LDA: Why Priors Matter, in: Proceedings of the 22nd International Conference on Neural Information Processing Systems, Curran Associates Inc., Vancouver, British Columbia, Canada. pp. 1973–1981.

Zuidema, P.A., Zagt, R.J., 2000. Using population matrices for long-lived species: a review of published models for 35 woody plants, in: Demography of exploited tree species in the Bolivian Amazon. PROMAB Scientific series 2, Bolivia, pp. 159–181.

Supplementary material

R script and model used in JAGS with an illustration of the model validation.

Fig. 8: Example of convergence for $\pi_{1\,2}$ (2014-2015).

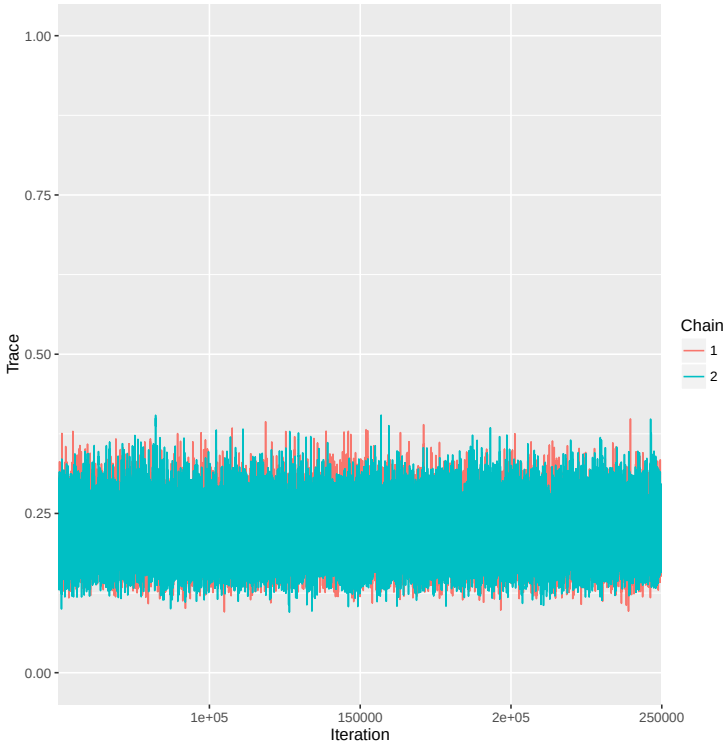


Fig. 9: Example of the obtained Gelman test for $\pi_{1\,2}$ (2014-2015).

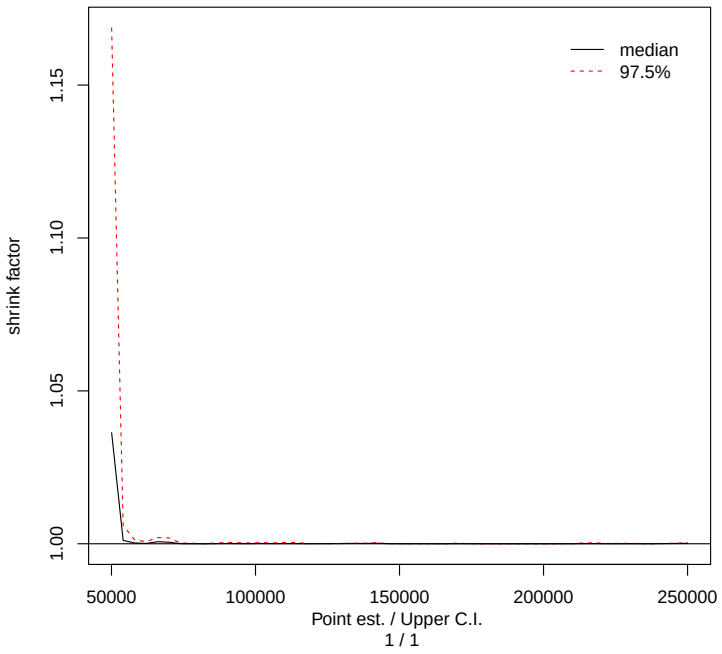


Table 7: Posterior densities of the transition survival probabilities at the level of the whole studied area, respectively, over 2013-2014 on the left and over 2014-2015 on the right: death is shown in red, transition in green, and idle in blue.

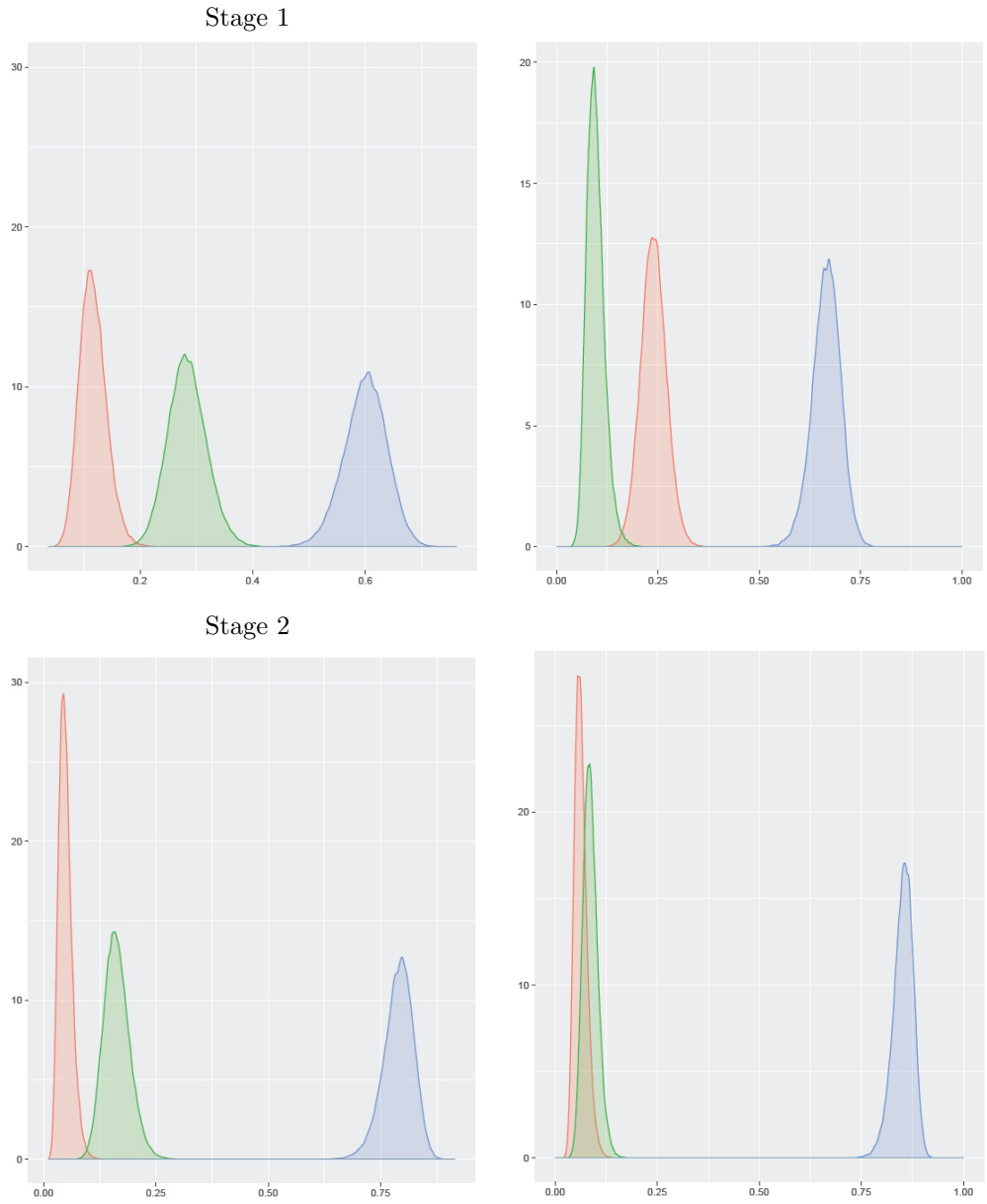
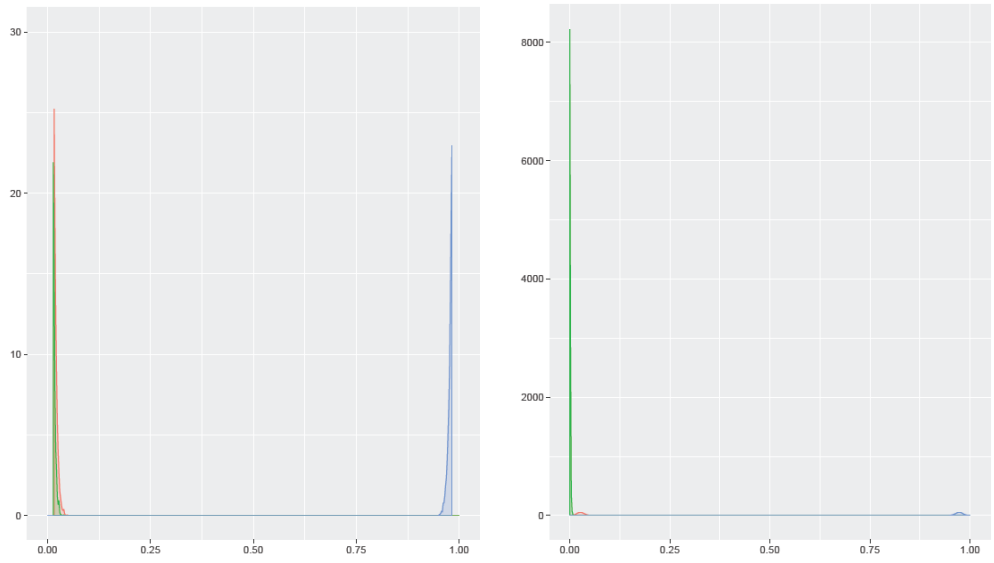


Table 8: Posterior densities of the transition survival probabilities at the level of the whole studied area, respectively, over 2013-2014 on the left and over 2014-2015 on the right: death is shown in red, transition in green, and idle in blue.

Stage 3



Stage 4

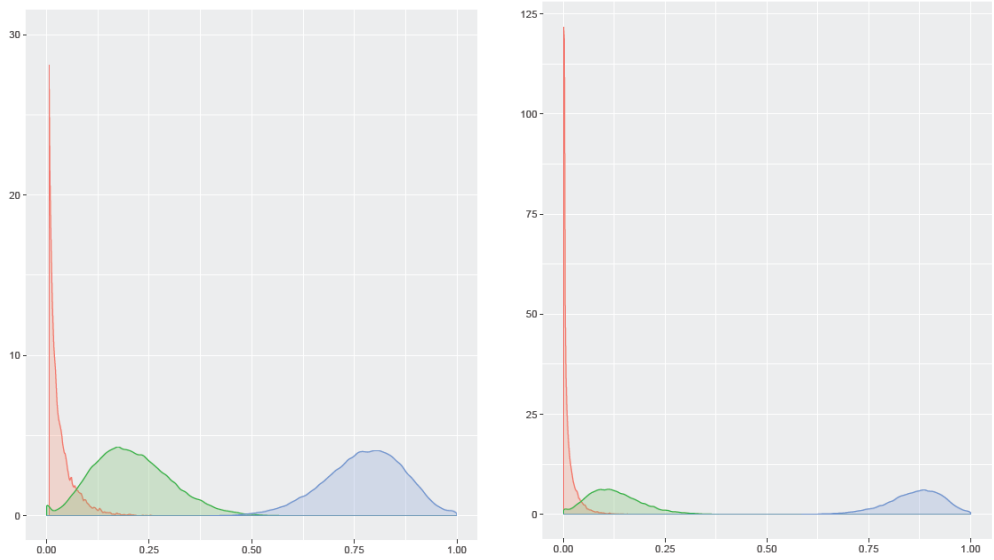
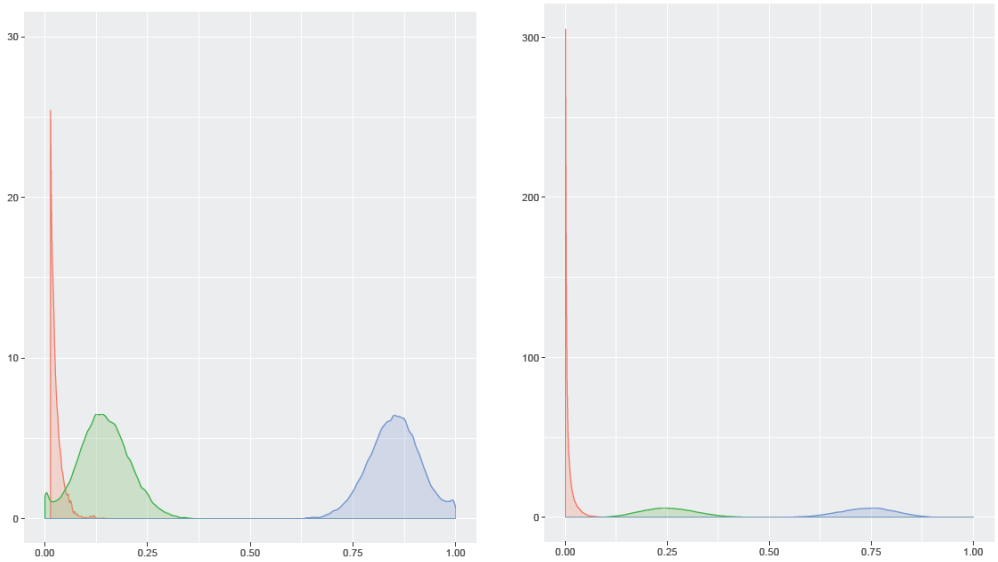


Table 9: Posterior densities of the transition survival probabilities at the level of the whole studied area, respectively, over 2013-2014 on the left and over 2014-2015 on the right: death is shown in red, transition in green, and idle in blue.

Stage 5



Stage 6

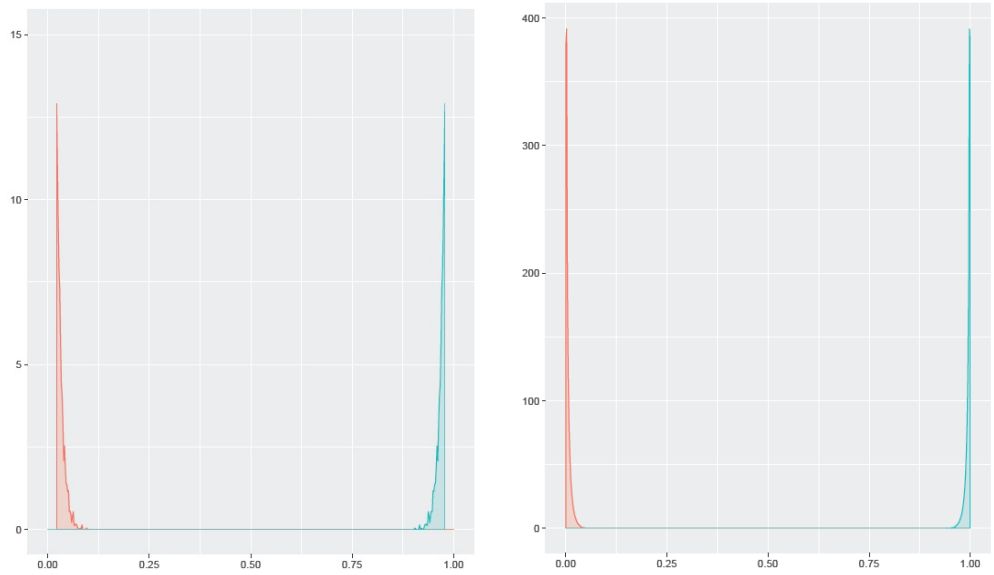


Fig. 10: Adjustment of the distribution of the number of fruits per palm tree.

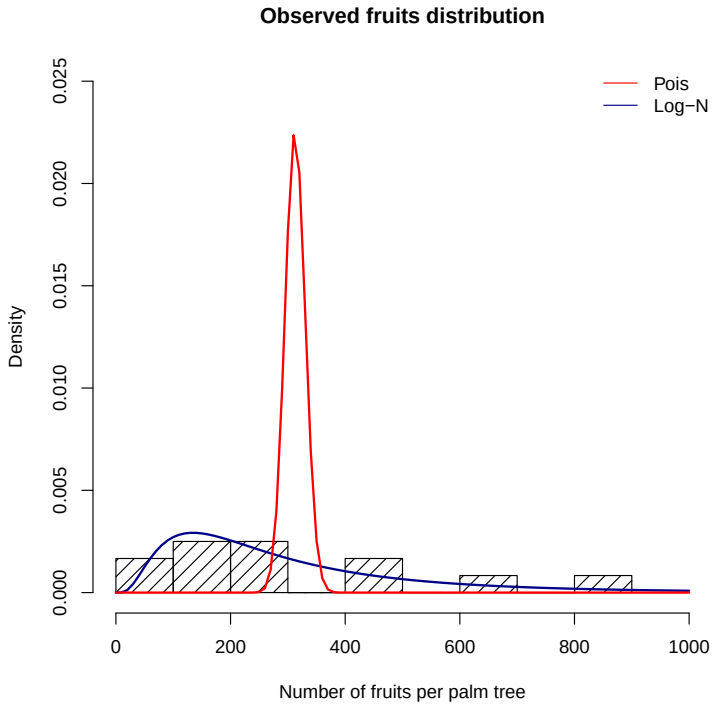


Fig. 11: Estimation of the probability of transition from stage 2 to stage 3 according to the coordinates of the pieces of land.