**Coffee monoculture trend in tropical agroforested landscapes of Western Ghats (India)**

***Cédric GAUCHEREL \*(1, 2), Julie ALET (1), Claude GARCIA (3, 4)***

**Supplementary material**

## Appendix S1. Kottoli village and the farm typology.

The project “Connecting Environmental Services and Market Values of Coffee Agroforestry” (CAFNET), funded by the European Union from 2007 to 2011, was designed to assess and link the provision of environmental services to the market value of the coffee produced under complex agroforestry systems (Ambinakudige and Choi, 2009). CAFNET specifically targeted 126 villages in the central watershed of the district for its research. One of such villages was Kottoli, with circa 2000 inhabitants distributed over 5 km². Its elevation ranges between 850 and 1000 metres, and a river crosses the north-western part of the village (Fig. S1).

We mapped the village using survey maps obtained from the Revenue Department, Google Earth® visual interpretation, and GPS groundtruthing. Terraced rice fields were classified into three land-cover types: cultivated field, fallow, and converted terrace. Half of the terraced fields were still cultivated, one-third left as fallow, and the rest converted to other crops. Wooded patches were classified according to (1) their nature: forest or agroforest, (2) the percentage of silver oak, and (3) their tree canopy cover (in percentages), and were mapped along 15 transects crossing the village. We pooled the habitations and water bodies with the major surrounding soil constituents (rice or forest), as they were small enough to not interfere in the land cover distribution. The wooded patches covered

68% and the terraces 32% of the total surface of Kottoli (Table S1). Coffee plantations with native tree cover were the dominant land cover (21%), while forest remnants (patches of forests not converted to coffee, either because of their being held sacred and community-controlled, or because the private owner had not made the conversion) covered 9%. Silver-oak-dominated coffee plantations covered 25% of the total village surface, more than one-third of the wooded area.

We used the Wilcoxon rank-sum test. The samples here were the projected coordinates of the polygon centres of each land cover on an axis linking Kottoli’s south-western and north-eastern parts. With the null hypothesis of a uniform distribution of the land cover, a low p-value (below 10-2) indicated a significant gradient of this land cover across the village. The main gradients observed in the 2011 Kottoli map were: a positive fallow paddy gradient (Wilcoxon p-value = 2.2 10-16), a positive silver oak coffee plantation gradient (1.14 10-5), and a weak and positive private forest gradient (1.9 10-3).



Figure S1. Location maps of Kottoli village. The 2007 land-cover map is based on IRS P6 LISS III images and was reclassified by manual digitization.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Initial – 1950**  | **Observation – 2011**  | **Simulation – 2010** |
| *km2* | *%* | *km2* | *%* | *km2* | *%* |
| 5 - Jungle coffee | 0 | 0 | 1 | 20 | 1.05 | 21 |
| 7 - Mix SO-jungle coffee | 0 | 0 | 0.83 | 17 | 0.83 | 17 |
| 1 - Cultivated paddy | 1.59 | 32 | 0.76 | 15 | 0.81 | 16 |
| 6 - Open jungle coffee | 0 | 0 | 0.71 | 14 | 0.71 | 14 |
| 2 - Fallow land | 0 | 0 | 0.58 | 12 | 0.6 | 12 |
| 4 - Private forest | 3.38 | 68 | 0.55 | 11 | 0.44 |  9 |
| 8 - SO coffee | 0 | 0 | 0.29 |  6 | 0.38 |  8 |
| 3 - Converted Paddy | 0 | 0 | 0.25 |  5 | 0.16 |  3 |
| 9 – Devarakadu | 0.03 |  1 | 0.03 |  1 | 0.03 |  1 |

Table S1. Detailed land-cover composition of Kottoli village at various dates.

Semi-structured interviews were used to examine the past and present management practices of the farmers in relation to their plantations and terraces, as well as their plans for the future (Olivier de Sardan, 2003; Bierschenk and Olivier de Sardan, 1997). We covered the village door-to-door in March and April 2012, and carried out in total 43 in-depth interviews in English or in Kannada (the local language). We asked 5 general questions and 2 questions about problems encountered in the farm, 16 questions about coffee plantations, 7 questions about paddy fields, 2 questions about the past, and 5 about perspectives. In addition, targeted questions on specific topics such as the conversion of terraces were put to the remaining 83 farmers of the village in subsequent field campaigns. These targeted questionnaires aimed at defining the driving rules of the landscape modelling. Based on these surveys, we created a typology of the farmers through Principal Component Analysis (PCA) (Bidogeza et al., 2009; Tittonell et al., 2010), first on the 43 in-depth interviews to reduce the number of significant variables, and then on the complete set of 126 interviews to feed the DYPAL model. The variables used in the PCA were chosen for their contribution to the overall variance, their relevance in the landscape modelling, and their frequency of appearance (Table S2). The age of the silver oak trees was the only variable providing temporal information. Proportions rather than absolute values for the area of converted and fallow terraces were used in the PCA, as farmers had an easier time relating to this as part of a strategy. Finally, a hierarchical ascendant classification (HAC) was realized to identify the dominant farm groups.

A first PCA based on the 43 detailed farmer interviews, and explaining 74.2% of the variance with three principal components (PCs), enabled a selection of the six most relevant variables (out of 16 available in total) for the farm typology: (1) area under coffee cultivation (*coff*), (2) area under rice cultivation (*padd*), (3) area of converted rice fields (*conv\_prop*), (4) area of rice fields left fallow (*fall\_prop*), (5) ages of the silver oak trees in the coffee estate (*age\_so*), and (6) total number of permanent workers (*perm*). A second PCA based on the 126 farmers explained 82.1% of the variance with three PCs. The first PC (49.8%) reflected the capital of the farm, combining the area under cultivation (*coff* and *padd*) with the number of permanent workers (*perm*). The second PC (18.7%) highlighted the effort invested in phasing out rice cultivation, as it was positively correlated to the area of converted rice fields and negatively correlated to the area left fallow. The third PC (13.6%) corresponded to the ages of the silver oak plantations.

The HAC classification distinguished four farm groups: G1) “Large coffee producers” (5): farmers owning extensive estates in Kottoli, from 6 to 16 ha of coffee plantations, and 2 to 8 ha of rice fields. These owners managed their estates essentially through their permanent workers, concentrated mainly on coffee production, invested in planting silver oak, and either converted their rice fields or left them fallow. G2) “Focused coffee producers” (12): farmers who owned relatively small rice and coffee estates (around two ha), and left their rice fields fallow. They had had silver oak for more than 30 years. G3) “Marginal coffee producers” (18): farmers owning 0.1 to 2 ha of coffee and up to 4 ha of rice. Almost all their rice fields had been converted into coffee, areca, or banana. These farmers in general had started planting silver oak less than 10 years ago. G4) “Subsistence farmers” (91): the most numerous and the poorest group in the sample. They had small holdings (1 ha of coffee and 0.4 ha of rice on average). Almost all their rice fields were under cultivation, and they had started only recently to plant silver oak. They had no permanent workers, and the only human capital available on their farms was the family.

The typology tuned to four homogeneous farmer groups was seen to be an efficient trade-off between an over-simplified simulation and a higher number of groups, leading to a difficult characterization of the driving rules for each group. We neglected social mobility and the possible changes of farm types over 60 years, although DYPAL would have been able to handle those aspects if the information had been available. Integrating demographic changes that alter the groups and their respective weights in the population opens up avenues for future research. Off-farm income and remittances also appeared as important additional variables that we had not been able to capture and integrate in the model. Our farmer typology map (Fig. S2d) did not feature the exact locations of the farms and fields attached to each individual farmer, as these were often unknown. The map was used instead to highlight the spatial distribution of the farmer groups. "Large coffee producers" (G1) and "Focused coffee producers" (G2) mostly belonged to the rich northern part of Kottoli, where the best soil was to be found. "Marginal coffee producers" (G3) and "Subsistence farmers" (G4), representing 72% of the farmers interviewed, were spread across the landscape.



Figure S2. Automata according to the farm groups (abscissae) and the events (ordinates) with the following land-cover codes: cultivated paddy **P**, fallow paddy **F**, converted paddy **C**, forest **F**, agroforest of jungle trees **J**, open jungle **OJ**, mixed **Mix,** and pure silver oak plantations **SO**. Grey arrows mean that the set of rules possibly apply to other farm types and over the whole time frame of the simulation.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Variables** |  **Nature of data**  | **Data information** |
| 1 | coffee surface |

|  |
| --- |
|  |
|

 | quantitative and non-binary  | capital |
| 2 | cultivated paddy surface |
| 3 | % converted paddy | practices |
| 4 | % fallow paddy |
| 5 | age of silver oak plantation |
| 6 | number of permanent employees | social status |
| # | irrigation use |  binary |   |
| # | forest surface |  uneven  availability |   |
| # | crop production |  uncertain |   |

Table S2. Selection of the variables deduced from the Kottoli interview for the PCA analysis.

## Appendix S2. The DYPAL model and its language

DYPAL is a platform to model landscape dynamics, already detailed in the literature (Gaucherel et al., 2010; 2012). The DYPAL models assume that landscapes are patchy, divided into landscape units (polygons, polylines, and points), with a main attribute reflecting its land cover. The core of the model is a set of driving rules aiming at changing the landscape through stochastic and non-Markovian events. In addition to these land-use changes, configuration (geometrical and/or topological) changes of landscape units can alter the shapes, the spatial arrangement, and the sizes of the patches, thus reflecting ecological processes such as fires (contraction), forest expansion (dilation), as well as man-mediated processes such as farm segmentation or patch mergers. The length and time step of each simulation are defined by the model user, and the main model outputs are land-cover maps for each time step as well as their associated metrics. DYPAL is freely available on the senior author’s website (<http://amap-collaboratif.cirad.fr/pages_chercheurs/index.php?page=softwares>).

The driving rules representing farmers’ decisions are summarized into automata (decision trees), in which land covers are symbolized by nodes and their transitions by arrows connecting the nodes (Fig. 2). Each rule is associated with a probability of occurrence, an acting period, and several conditional constraints (e.g. a neighbourhood condition) called “rule parameters”. It was recently proposed to formalize these rules using a formal grammar inspired by DS² dynamical systems (Godin, 2000; Gaucherel et al., 2012). Under this convention, a rule is written as M(lc) 🡪 M(lc’), M being a module representing a landscape patch, lc and lc’ being two different states of its land-cover attribute. Conditions (*cond*) and probabilities (p0) may be used as follows, M(lc): *cond*: p > p0 🡪 M(lc’). Conditions are necessary but not sufficient; if the condition *cond* is fulfilled, the transformation will occur only with the p0 probability. Additional conditions (called “contexts”) can be added. For example, should we want to specify that a given transformation will happen only in farms using a particular production system, we would write:

F(f) < M(lc) 🡪 M(lc’) where the farm F has the production system f. If we wanted to specify that the patch needs a specific neighbour for the transformation to happen, we would write M’(lc’’) ≈ M(lc) 🡪 M(lc’) where the neighbouring patch M’ has the land cover lc’’. In the present instance, all the information on conditions, probability, and patch transformations was derived from the interviews, while that on historical transformations was provided by interviews and by the survey maps.

To finalize the modelling of the landscape of Kottoli, we needed three additional maps (Fig. S2). The 2011 map of Kottoli was drawn based on field surveys and groundtruthing as explained before. We used this map to validate the model by an accurate local and multiscale comparison, as explained in the next section (Fig. S2a). The 1950 map of Kottoli, featuring a mosaic of terraces and private forest patches (the oldest coffee plantations were 59 years old), was reconstructed from the 2011 map and interviews, as no other information was available apart from the poor-quality 1917 survey map. We are conscious that this map is an approximation only, but several simulation tests indicated that our results were not sensitive to the exact form of this map. We identified two hydromorphic zones: the river banks, defined as a 300m buffer from the river bed based on the farmers’ indications, and the dry zones. Finally, we built a farm layer on the basis of a statistical analysis of farmer interviews (Fig. S2e). The farm typology map links together farm polygons related to one of the four farm groups extracted from the PCA analysis. A simple algorithm based on a minimal distance association was developed to locate the farmers, their farm group, and the area of coffee and terraces they possessed.

A 5-year time step was chosen, and our simulation needed 12 time steps to cover the time span from 1950 to 2010. We found this time step to be a good compromise between detail (a time span short enough to allow a modelling of comparatively minute transitions, a task that would have been unwieldy if forcing down probabilities had been resorted to) and explanation power (a larger time span clubbing together several processes, thereby making the end result difficult to understand). Several conditions were applied to the Kottoli modelling: (1) multilevel condition: a rule applies if the patch belongs to a particular (farm or soil) type; (2) time validity: the rule applies for a specific duration; and (3) neighbourhood relations: the rule applies only if the patch is contiguous with other specific land covers. The two latter conditions only have already been used in earlier simulations (Castellazzi et al., 2007; Gaucherel et al., 2012). Composition changes (rotations) were developed for Kottoli dynamics. A trial-and-error calibration was performed on the basis of some composition indices (to quantify how close simulated compositions were to the observed composition) in order to define probabilities associated with each rule. This stage was helped by a detailed elasticity test highlighting the most sensitive driving rules to be finely tuned. We checked that simulations were stable for slightly different probabilities. Finally, we tested scenario simulations of increasing complexity to mimic more and more realistic driving rules and dynamics (the last one being named Hf).

The language (mathematical) formalization is an important result in itself. As an illustration, three contrasting driving rules out of 18 (Table 1) are set out here.

Rule n°1. M(lc) : (lc = 4) : p = 0.18 🡪 M(5)

This rule models the conversion of private forests (lc = 4) into jungle coffee plantations (lc = 5). It applies over 18% randomly chosen private forest patches at each time step of the simulations. Each rule is fixed during a simulation, as are their probabilities and conditions. Only the rule implementation is allowed to vary, depending on the period and events involved in each scenario. All rules were applied to the previously fixed time step state.

Rule n°5. F(2) < M(lc) : (lc = 6) & (D > 4) : p = 0.4 🡪 M(7)

This rule applies to open jungle coffee patches (lc=6) belonging to the farm group number 2 (F(2)), and starts at time step n°5 (D > 4, *i.e.* 1970). Only 40% of these randomly chosen patches will turn into Mixed Jungle-Silver oak coffee (lc = 7).

Rule n°11. F(1) & H(0) < M(lc) ≈ M(2|4|5|6|7|8|9) : (lc = 1) & (D > 9) : p = 0.1🡪 M(2)

The H(0) condition (hydromorphic layer, class 0: paddy can be converted) imposes the condition that the randomly selected rice field (lc = 1) patch should be distant enough from the river. In the event that this patch is contiguous with either converted rice fields or forested plots (≈ M(2|4|5|6|7|8|9), it can be converted (lc = 2). This rule starts at the step 10, and has 10% chance to affect any eligible cultivated rice field. The final simulated Kottoli map (Fig. 2f) has been computed on the basis of these rules (Fig. S2) and the initial available maps (Fig. 2a-e).

### Appendix S3. The MHM and CMP spatial analyses

The MHM and CMP methods have been extensively described and used elsewhere (Gaucherel et al., 2007; Gaucherel et al., 2008). These multiscale spatialized methods are both based on a circular moving window simultaneously crossing images to be analysed (MHM) and compared (CMP), combined with specific textural or similarity indices. They lead to local and quantitative analyses, which become multiscale when the size of the moving window varies (here from 10 to 150 pixels wide, with a 20 × 20m pixel size). Monoscale maps were finally averaged over all scales, pixel by pixel, into a multiscale image to quantify the spatial variations of the texture or similarity indices, while monoscale maps were averaged over space, image by image, into a profile to quantify the scaling variations of the texture or the similarities. All spatial analyses described in this study were computed ignoring the landscape's background and without border effects, with the Java MHM and CMP software (available on the senior author’s website (<http://amap-collaboratif.cirad.fr/pages_chercheurs/index.php?page=softwares>)).

In this study, we first computed land-cover connectivity (contagion) maps of Kottoli simulations and observations with the MHM software (Gaucherel et al., 2007). We used connectivity as a synthetic index, considering that land-cover maps would not match due to the stochasticity of each landscape simulation. Rather, multiscale connectivity computation provides an overall and statistical quantification of the landscape configuration, without favouring that of the land cover. Another advantage is that this index is highly robust and almost independent of the stochastic simulation run (less than 5% variations), thus allowing one to compute differences on a randomly chosen simulation of H0 and Hf simulations (Table S2). Simulated connectivity maps were then compared to the 2011 observed map on the basis of two complementary indices (Gaucherel et al., 2008): i) a distance index, computed as mean differences for absolute differences, and ii) Pearson’s cross-correlation for relative variations.

Once simulations were validated, and in order to understand the differential success of various land covers, we computed land-cover multiscale density maps with the MHM software, and compared them to the corresponding observed density maps with the CMP software. In this study, the density index is defined as the surface proportion of the analysed land cover within each moving window, while the chosen similarity index is the Pearson coefficient to highlight relative differences. We compared the forest, silver oak, and fallow correlation profiles to detect whether these land-cover distributions, at least, were correctly modelled. As an illustration, we show here the various correlation maps (Fig. S3b) and profiles (Fig. S3c) between Kottoli simulations and observations for i) the fallow rice field, ii) the silver oak coffee plantation and iii) the forest variable.

****

Figure S3. Cross-correlations between simulated and observed landscapes of Kottoli in 2012. The three land-cover density maps computed with MHM for observed (a, large) and simulated values (a, small) are compared to their respective CMP correlation maps (b) and profiles (c). Gray levels indicate density intensity between 0 and 0.07 (a), the colour scale indicates the cross-correlation intensity, ranging from – 1 to + 1 (b). Correlation profiles (c) computed for spatial scales ranging from 10to 150, are displayed forfallows (green), silver oak (blue) and forest (black) with dashed red segments highlighting some of their features (see the text).