ASSESSMENT TOOL / MODEL The potential of Complexity Index to analyze time-variant participatory approaches together with time-invariant household surveys

Dharani Dhar Burra[1] & Sabine Douxchamps[2]
[1]Principal Data Scientist / VARDA AG
[2]Forage Based Crop-Livestock Systems Specialist / The Alliance of Bioversity and CIAT

This study gives insights in how to compress historical patterns of land use (i.e. state-sequence), consisting of each crop/crop combination (as a state), and the recall period (as a sequence) into a single metric, to arrive at a Complexity Index (CI). The linkages between CI, performance, plot and household level characteristics, and assessment of within-site variation are addressed for a case study in the Central Highlands of Vietnam.

NOTE An example of this tool in use is included as a case study at the end of this learning material.

Overview

The Complexity Index (CI) is a sequence dissimilarity measure to quantify land use transition in an agricultural system. It is a composite measure that combines the number of transitions occurring on each plot across multiple land uses (i.e. states), across time (i.e. sequence) with the longitudinal entropy. The entropy is a measure of the diversity of states at site level at a given position in the sequence.2. The CI can be used in agriculture to better understand diverse ranges of determinants (specifically farm-performance indicators, plot-level, and household-level characteristics) of land use change (wherein each land use type is considered as a "state") across time. The process contributes to the identification of trade-offs (e.g. between crops that contribute to different objectives) and synergies which are relevant for sustainable land-use planning. 3. The CI reflects land use changes at plot level. When mapping the results in combination with their determinants, it highlights trends and tradeoffs at landscape level.

Relevance of the level of analysis

Complexity Index (CI) was originally designed to study sequence states, wherein the position of each successive state in a sequence receives a meaningful interpretation, either based on time taken or distance from the beginning of the sequence. Its use has been largely restricted to the analysis of longitudinal data, such as data describing careers or family trajectories. However, it can also be applied to non-temporal data such as DNA sequences.

It has been recently used for the first time in agricultural systems to better understand a diverse range of determinants (specifically farm-performance indicators, plot-level, and household-level characteristics) of land use change (wherein each land use type is considered as a "state") across time.

Model/tool description

What is the Complexity Index?

Transition processes are complex. For example, the transition from subsistence farming to more cash crops is not a unidirectional pathway, but rather a dynamic process. By understanding how changes contribute to performance outcomes (e.g. operating profit, labor dynamics, food availability) trade-offs and synergies can be identified and acted upon. To provide an example – Through the use of CI and their determinants, we see that farmers in a specific location lose crop diversity across space and time but gain in household incomes, while those that maintain crop diversity across space and time have comparatively lower incomes and poorer socio-economic conditions. Here using CI and determinants, see spatial and temporally disaggregated tradeoffs between crop diversity and household incomes

The CI is based on the number of transitions occurring on a single plot across multiple land use types over time, and longitudinal entropy. The value of longitudinal entropy is 0 if the same land use type is identified across time (tree-based systems), and the value is maximum if all land use types assessed are present in equal proportion across time (e.g., alternating between seasonal crops). The CI enables compression and quantification of land use complexity across time. It can be used in combination with time-invariant household and plot-level surveys, to determine drivers of transition from sustenance-based farming to market-driven intensive farming systems.

Although, in consequence, the transformation impedes the study of spatial-temporal land use patterns, it allows making linkages with indicators that are typically not captured over time. CI calculation is simple, with a low level of parameter uncertainty, and can be applied at a variety of levels. Indeed, the level of definition of transition depends on the user: for example, if the inclusion of hedgerows in a field qualifies for a new land use type, the resulting CI will be very different than if these types of land use are considered equal. This makes CI very flexible in its use and able to capture a wide variety of transition types, but also might restrict its potential for meta-analysis if the level of definition varies significantly between studies.

When CI is mapped, it highlights trends at the landscape level. Similarly, to the definition of land use, the size of the unit of observation will influence the type and scale of the drivers of change observed. Compared to land use changes observed based on satellite data, the level of definition is much more precise, and more appropriate to capture crop sequences and seasonal changes. Developing context-specific CI thresholds associated with system quality indicators would be useful as ecological engineering control to inform decision makers on the pace of agricultural transformation and its environmental impacts. Aggregated Complexity Index. CI reflects land use change at the plot level, which enables a relationship with plot-level surveys. However, since we also had time-invariant household surveys in addition to plot-level surveys, and each household manages multiple plots, we developed a new metric called aggregated Complexity Index (agg-CI), which averages CI across these multiple plots belonging to a single household. In this study, performance indicators and household characteristics are defined at one point in time (i.e. in 2017 and hence time-invariant) whereas agg-CI is the aggregate result of 10 years.

This assumes that there has been no major alteration in the households and that the present situation comprehensively summarizes all happenings on the farm during the last ten years.

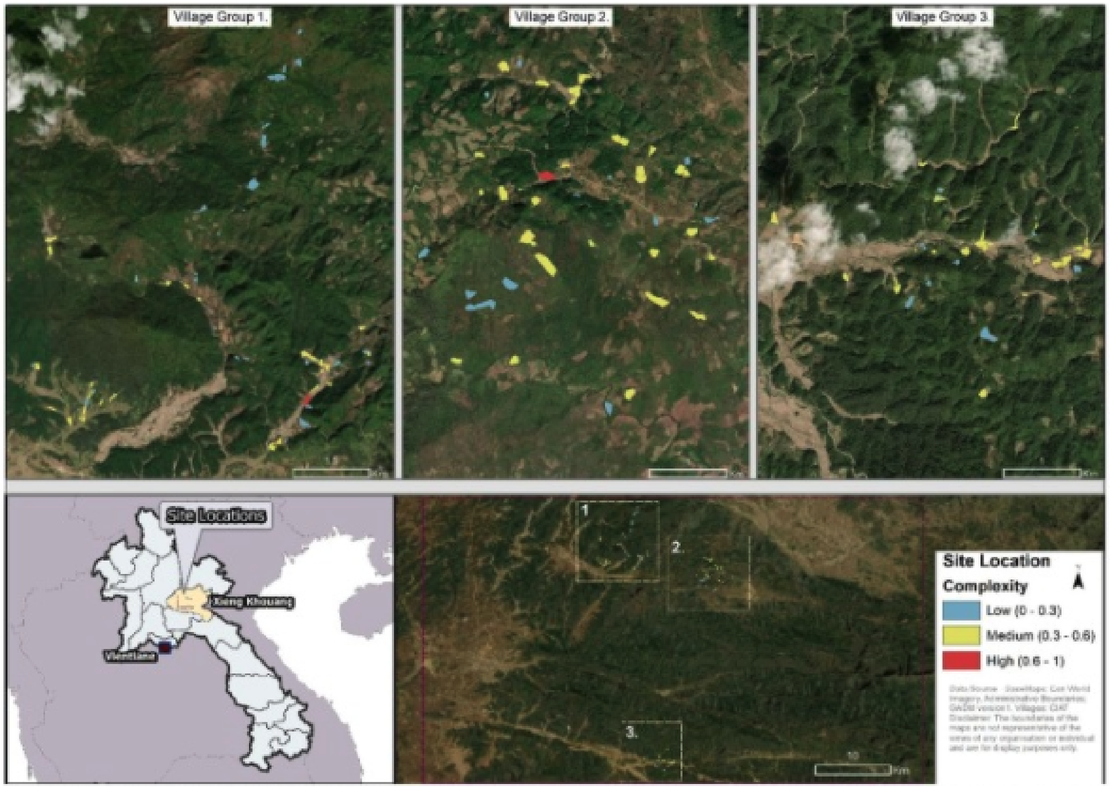

Figure 1. Spatial distribution of the complexity index (CI) in Xieng Khouang, Laos. Graphic from Burra et al., 2021.

What are the disadvantages of the Complexity Index?

The disadvantage is that CI integrates seasonal information without consideration for the novelty of the crops: rice-fallow rotations are treated equivalently to switching from one crop to a completely new one (from coffee to sugarcane, for example). This has no consequence on the complexity of the system, but care must be taken in not translating this complexity as a measure of diversity, which is best considered by looking at

Future research directions

Future studies could bring more time variant perspective into household and plot level characteristics if some form of agricultural census data is available, but it would be difficult to do so on a recall basis, as farmers might have difficulties remembering some of this information precisely for remote years. Still, some form of aggregation would need to take place for a regression with CI. The use of plot level and agg-CI in this context is unique, and the models obtained are parsimonious and provide a significantly better fit than the full model. Regression analysis suggests that, in order to obtain robust results, there is a need to calculate CI scores from a larger set of households and their plots, to obtain significantly higher resolution between households, plots and their locations based on the derived CI scores.

Details for potential users

Proposed users - Scientists with knowledge of R
Key actors/stakeholders/beneficiaries - Scientists, land-use planners
Model input - Time series land use data
Model output - A single metric called complexity index (CI) that combines time series land use data into a single metric

Key terms

- **Complexity Index** - combines the number of transitions occurring on each plot across multiple land uses (i.e. states), across time (i.e. sequence) with the longitudinal entropy. CI is calculated using the following equation:

$$C(x) = \sqrt{\frac{\ell_d(x)}{\ell(x)} \frac{h(x)}{h_{max}}}$$

  wherein C(x) is the complexity index (CI) of a given plot x, hmax is the theoretical maximum value of the entropy given the state i.e. hmax = log (a). The entropy is a measure of the diversity of states at site level at a given position in the sequence.

- **Transition rate** - rate of probability of transition between all combinations of plot level land use types captured across time, thereby providing a proxy for stability of a specific land use type. The transition rate between two states si and sj is calculated using the following formula:

$$\frac{\sum_{t=1}^{L-1} n_{t,t+1}(s_i, s_j)}{\sum_{t=1}^{L-1} n_t(s_i)}$$

  wherein p(sj|si) is the probability of switching at a given position from state si to sj, L is the maximum observed sequence length, nt(si) is the number of sequences that do not end in t with state si at position t, and nt,t+1(si,sj) is the number of sequences with state si at position t, and state sj at position (t + 1).

Manuals, tutorials, or other learning materials

State sequence in R:
https://cran.r-project.org/web/packages/TraMineR/vignettes/TraMineR-state-sequence.pdf

Key references

D. Burra, L. Parker, N. T. Than, P. Phengsavanh, C. T. M. Long, R.S. Ritzema, F. Sagemueller, S. Douxchamps (2021). Drivers of land use complexity along an agricultural transition gradient in Southeast Asia. Ecological Indicators 124, 107402 (DOI 10.1016/j.ecolind.2021.107402)
https://www.sciencedirect.com/science/article/pii/S1470160X21000674

Case Study

In the Central Highlands of Laos and Vietnam households with higher food availability are half as likely to transition, while in the Lao uplands, land use complexity was significantly correlated with the Progress out of Poverty index.

Time period (or an indication): 2007-2017
Key actors/stakeholders/beneficiaries: smallholder crop-livestock farmers of the uplands

Applying the model: The Complexity Index has been tested across two sites in Southeast Asia.
In Vietnam's Central Highlands (CH), relatively well-educated migrants have increasingly applied intensive agricultural practices in market-oriented intercropped tree-based systems during the 2006–2017 period, accompanied by a decrease in monoculture production and fallow land. These 'highly transitioned' systems result in relatively high income and food availability. In the Xieng Khouang (XK) plateau of Laos, self-sufficient farmers are in the initial stages of reducing forest and fallow land to increase their rain-fed low-input staple crop production, with progressive inclusion of vegetable crops and forages as well as several cash crops.

Land use dynamics vary strongly between the two sites, with 66% of the land use types in site XK being completely replaced by others during the recall periods, compared to only 15% in site CH. Associated key drivers of change also differed significantly between the two sites: while end use of the agricultural products is the main driver behind land use changes at site CH, the relationship is more complex at site XK, with changes associated with topography and management. Households with higher food availability in site CH are less likely to show transition (i.e., lower aggregated complexity index), while in site XK, aggregated complexity index was significantly correlated with Progress out of Poverty index.