

I. Some Helpful General Definitions

Alphabet – a finite set of symbols.

Full shift spaces (full shifts) – if 'A' is a finite alphabet then the full shift space (full A-shift) is the collection of all bi-infinite sequences of symbols from 'A'. Each specific symbol sequence is called a point of the full shift.

LANDIS-II – a spatially explicit and spatially interactive forest landscape modeling environment. Formally, LANDIS-II is a process-based simulation model that generates spatio-temporal trends (patterns) of vegetative (forest) change based on (1) ecological mechanisms and processes; (2) the effects of a heterogeneous environment; and (3) the effects of disturbance on the landscape

m-order Markov chain (bigram model) – a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event; a statistical language model that uses a two symbol sliding window to assign 1-step conditional probabilities

Patch – in an agent-based model, this typically refers to a pixel in the 2D grid that defines the environment of the model. In LANDIS-II this term has the additional meaning of a collection of species cohorts that occupy a pixel, or more generally, a discrete area with distinguishable boundaries that has relatively homogeneous environmental conditions and biotic components

Phase space – the n-dimensional space generated by the evolution of the n-vector of a system's state variables through time. The evolution of a dynamical system can be seen a transformation that maps phase space onto itself: the system traces various trajectories in phase space over time

Point – a single, bi-infinite sequence of symbols from an alphabet

Shift space (shifts) – the subset of points in a full shift satisfying a fixed set of constraints

Symbolic dynamics – the course-graining approach to the analysis of dynamical systems where the behavior of the system is tracked as a sequence of symbols generated by discretizing time and phase space. The idea is to (1) divide up the set of possible states of a system into discrete regions; (2) identify each region with a specific 'symbol'; and (3) track the evolution of the system in discrete time as it moves through different states by writing down the sequence of symbols. The dynamics of the system are encoded, depending on the grain, in a (potentially infinite) sequence of symbols

Vertex shifts – A Markov chain can be mapped to a directed graph, $G = \{V, E\}$, with a set of vertices and edges. The vertices can be defined by an alphabet. Edges have probabilities conditional on the originating vertex and represent the probability of generating a bi-gram (two symbol sequence). Sequences or points are (potentially infinite) paths on G . Vertex shifts specify only the allowed sequence of outcomes; that is, only those bi-grams with a probability greater than zero (or some fixed value)

II. A Few Definitions of Sustainable Agriculture

The US Congress idea – [7 U.S.C Sec. 3103.] Definitions.

- (19) The term sustainable agriculture means an integrated system of plant and animal production practices having a site-specific application that will, over the long-term –
- (A) satisfy human food and fiber needs;
 - (B) enhance environmental quality and the natural resource base upon which the agriculture economy depends;
 - (C) make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls;
 - (D) sustain the economic viability of farm operations; and
 - (E) enhance the quality of life for farmers and society as a whole.

The USDA idea – Sustainable agriculture is "a way of practicing agriculture which seeks to optimize skills and technology to achieve long-term stability of the agricultural enterprise, environmental protection, and consumer safety. It is achieved through management strategies which help the producer select hybrids and varieties, soil conserving cultural practices, soil fertility programs, and pest management programs. The goal of sustainable agriculture is to minimize adverse impacts to the immediate and off-farm environments while providing a sustained level of production and profit. Sound resource conservation is an integral part of the means to achieve sustainable agriculture." [United States Department of Agriculture (USDA) Natural Resource Conservation Service (NRCS) General Manual (180-GM, Part 407)]

The EU idea Sustainable agriculture aims to: (1) Produce safe and healthy food; (2) Conserve natural resources; (3) Ensure economic viability of farmlands; (4) Deliver services for the ecosystem and for people (biodiversity, soil conservation, nutrient storage, carbon storage); (5) Manage the countryside (preserve habitats and landscape beauty); (6) Improve quality of life in farming areas; and (7) Ensure animal welfare. ["Sustainable agriculture for the future we want"; EU Commissions on Development and Cooperation and Agriculture and Rural Development, 2012]

The Union of Concerned Scientists idea – "Sustainable agriculture does not mean a return to either the low yields or poor farmers that characterized the 19th century. Rather, sustainability builds on current agricultural achievements, adopting a sophisticated approach that can maintain high yields and farm profits without undermining the resources on which agriculture depends." ["Sustainable Agriculture – A New Vision"; Union of Concerned Scientists, 1999]

III. Options for Key Monitoring Variables

There are many potential sources for key monitoring variables (KMTVs) and unfortunately, the selection of KMTVs is an open-ended problem of logical analysis. Measurement scales for KMTVs may be nominal, ordinal, or ratio. Once the m-dimensional vector of KMTVs has been defined, the n-dimensional vector of required empirical variables (REVs) becomes fixed: REVs are the directly measured attributes (i.e., the system signal outputs) that are either (1) KMTVs themselves, or (2) the conditional variables necessary for the derivation of other KMTVs. REVs can be obtained from previous observational studies, current observational studies, or simulation studies.

Source 1 – Mayan *milpa* practices (Classical Period, 300 – 900 AD)

1. rate, magnitude, and extent of agricultural disturbance
2. land use/land cover ratio
3. biodiversity ratio(s)
4. human density
5. soil density of arbuscular mycorrhizal fungi
6. soil density of ectomycorrhizae
7. soil density of N-fixing bacteria

Source 2 – Landscape ecology (structural pattern metrics)

1. composition
 - (a) proportion of each patch type
 - (b) Shannon diversity index
 - (c) Shannon evenness index
2. configuration
 - (a) lacunarity
 - (b) relative contagion index
 - (c) fractal dimension

- (d) patch edge-to-patch area ratio
- (e) correlogram (spatial autocorrelation)
- 3. texture gradients
 - (a) surface roughness
 - (b) shape of the surface height distribution
 - (c) angular texture
 - (d) radial texture

Source 3 – Traditional ecosystem ecology (processes and functions)

1. net primary productivity
2. mass of organic (reduced) soil carbon
3. water use efficiency
4. standing biomass
5. water holding capacity / leaching potential

Source 4 – MaxEnt ecology (John Harte, UC Berkeley)

1. species abundance distribution
2. species area relationship
3. endemics area relationship
4. community energy abundance relationship
5. cross-community scaling relationship (a mass – abundance relationship)
6. link distribution in a species network

IV. Options for Feature Extraction

Feature extraction, an absolutely key step in the proposed methodology, means examining the entire set of KMVVs (model output) and defining a small set of unique states using some form of discriminant analyses or clustering algorithm. The required outcomes of this discrimination process are: (1) a unique prototype or reproduction vector for each morph (state); (2) a disjoint and complete partitioning of the entire m-dimensional KMVV space into mutually exclusive and exhaustive subspaces which will define the 'morphs'; (3) a unique, m-dimensional joint probability distribution of KMVVs for each morph as an empirically-derived multi-dimensional histogram.

Feature extraction is simply the task of separating groups of like objects but is, in fact, a substantially non-trivial task in m -dimensions. Some useful tools for this task include,

- hidden Markov models (HMMs)
- multi-section vector quantizers
- support vector machines
- k-means nonhierarchical clustering algorithms
- principal components analysis
- hierarchical neural networks

Note that evaluating the goodness (accuracy) of any produced feature subset is dependent on what measure is used. Options include various information measures, distance measures, dependence measures, consistency measures, and accuracy measures.