# Bayesian Multivariate Markov Processes for a Network Flow Optimization Problem

P. Vanegas, K. Abayomi

***Abstract*—Vanegas [2009] considered the reallocation of land areas after land development as an Integer Programming (IP) control problem where the objective function is the minimization of sediment flow to riverbeds subject to physical-hydrological constraints via a Network-Flow (NF) formulation. The control variables in Vanegas [2009] are calculated from spatial data via raster maps via affine (piecewise-linear) functions; the optimal locations are discovered by Heuristic search.**

**This paper updates this setup with three key advances: (1) the spatial data via raster map is regarded as probabilistic having non-degenerate *prior* distributions in a Bayesian formulation; (2) the heuristic search is replaced by a Markov Chain progression where the control variables - still calculated via affine transformations - have their (*posterior*) probability distributions updated; (3) the heuristic search (for the optimal locations) is supplanted by the Markovian progression.**

**Then, for an initial set of raster maps, this approach yields a probability distribution for each spatial location, i.e. which locations are more probably optimal to reforest.**

## INTRODUCTION

We reframe the optimization problem in Vanegas [2009] in three important ways:

1) Put the raster map data (matrices) as random variables with non-degenerate probability distributions.
2) Substitute the objective function in the LP with posterior distributions for the (random) locations to minimize sediment flow.
3) Replace the heuristic search with a probabilistic version via Markovian conditional probabilities on the affine transformations of the data for the control variables

We can consider this - as in the original paper - as a version of a Graphical Model if even just for illustrative purposes. What this reframe does is replace a Directed Graph (DG) in the original frame with a Probabilistic Graphical Model (PGM) in the new version [Koller, Friedman 2009].

The approach is Markovian in the sense that we (necessarily) place conditional dependency restrictions on the - now - random data and dependent variables.

This allows us to exploit the Chapman-Kolmorogov (CK) equations (discussed below), the Copula equivalence with the CK equations (below as well), and places restrictions upon the Probabilistic Graphical model which are necessary for estimation.[1]

By reframing this optimization problem as a Multivariate Markov Process we are, directly:

- Replacing the objective function Equation

[1]The simplest graphical models are non-Acyclic, i.e. trees. We are close to that specification here.

(6) in Vanegas [2009] - minimizing Effective Accumulation (EA) at the 'root' or outlet node of the graph - with a *posterior* probability distribution for EA.

- Replacing the deterministic relationships between the raster (matrix) data and the control variables with, necessarily, probabilistic ones.

So what this scheme will generate is not a fixed set of nodes for deforestation for a particular set of initial conditions/raster data but maps with continuous probability of *posterior likelihood* of being reforested.

## I. Random Data

The LP setup in Vanegas [2009] specifies 10 matrices as input data to the LP scheme. Our first step is to outline and comment on each of these. We relabel them here as $\underset{\sim}{X}_k$ where $k$ is a alphanumeric index. For each we assign a (possible) *prior distribution*. In the specification below we consider the cells in the Vanegas setup probabilistically independent: this means the elements of each matrix $\underset{\sim}{X}_k$ are independent. This will be re-addressed in the complete probabilistic scheme - as there of course is a functional relationship between each matrix entry/graph node.

Lastly, as the raster maps - as data - may be easily standardized, we choose distributions for these random variables with minimal support. Here 'is distributed as' is written $\sim$:

- Flow Direction: We let each element in

$$\underset{\sim}{X}_{FD} \sim DiscUnif(0, 8) \tag{1}$$

that is a discrete uniform random variable with equal mass on each of the possible flow directions.

- Flow Production: Each

$$\underset{\sim}{X}_{FP,0}, \underset{\sim}{X}_{FP,1} \sim N(0, 1) \tag{2}$$

without loss of generality.

- Flow Factor: Each

$$\underset{\sim}{X}_{FF,0}, \underset{\sim}{X}_{FF,1} \sim Unif(0, 1) \tag{3}$$

without loss of generality.

- Breakpoint 1/Retention Capacity: Each

$$\underset{\sim}{X}_{BP1,0}, \underset{\sim}{X}_{BP1,1} \sim Unif(0, 1) \tag{4}$$

without loss of generality.

- Breakpoint 2: Each

$$\underset{\sim}{X}_{BP2,0}, \underset{\sim}{X}_{BP2,1} \sim Unif(0, 1) \tag{5}$$

without loss of generality

- Streams: Each

$$\underset{\sim}{X}_{ST} \sim Ber(0, 1) \tag{6}$$

These matrices are the independent variables, if you will, of the LP in Vanegas, in that they do not functionally or probabilistically depend on other data or random variables. Further - the *initial values* are the only real observed values: the remainder of the control variables in the Vanegas LP are calculated directly from these.

The parameterizations of the prior distributions here are *non-informative*: let $\gamma \in \Gamma$ be the collection of parameters and let $\pi(\gamma) = 1$, i.e. uniform over the support, for all $\gamma \in \Gamma$.

Lastly here I comment that this approach

is *fully Bayesian*: what this setup generates, for a instantiation of data (set of raster maps), are *posterior distributions* for the control variables. Contrast this with the original method in Vanegas [2009]: there a particular set of raster maps - data - yields *deterministically* a set of locations for reforestation. This *probabilistic* approach generates a probability distribution for reforestation at each node/location, as well as for EA and the other variables of interest.

## II. Dependent Variables

The flow delivery among nodes in Vanegas is governed by a piecewise linear function $g$; we will regard this as

$$\mathbf{Y} = \mathbf{g}(\mathbf{X}) \tag{7}$$

Here the bold $\mathbf{g}$ indicates that on the entire graph the transformation function is vector valued.

From first principles the probability distribution of $\mathbf{Y}$ will be

$$\mathbf{Y} \sim f(g^{-1}(\mathbf{X}))|\mathbf{J}| \tag{8}$$

where $|\mathbf{J}|$ is the determinant of:

$$\mathbf{J} = ((\partial \mathbf{g}^{-1}(\mathbf{Y})/\partial \mathbf{Y}))_{ij} \tag{9}$$

the Jacobian (or the matrix of partial derivatives of the transformation, component wise by the matrix and vector valued elements of $\mathbf{Y}$ and $\mathbf{g}$).

Now we can exploit two things in Vanegas 2009. First that the function $\mathbf{g}$ is linear. This means that

$$|\mathbf{J}| = \mathbf{c}^T\mathbf{X} \tag{10}$$

the determinant of the Jacobian will be a linear combination $\mathbf{c}$ of the instantiated values of the random variables of $\mathbf{X}$.

Second, as in Vanegas, let

$$A_{ij} = \mathbb{1}_{\{node_{ij}\ is\ reforested\}} \tag{11}$$

where $\mathbb{1}_{\{\}}$ is the indicator function. So $\mathbf{S}$ is the matrix collection of which nodes are reforested. Now let

$$\underset{\sim}{X}^T_{m,\cdot} = [\underset{\sim}{X}_{m,0}\ \underset{\sim}{X}_{m,1}] \tag{12}$$

be the matrix-wise concatenation of the raster maps for un-reforested and reforested locations for variable $m \in \mathcal{M}$, where $\mathcal{M}$ is the list of variables.

Put

$$\mathbf{X}^*_m = \underset{\sim}{X}^T_{m,\cdot} \otimes \mathbf{A} \tag{13}$$

so $\mathbf{X}^*$ is the matrix wise concatenation of all variables. Then the *prior distribution* for the dependent variables will be

$$\mathbf{Y} \sim f(g^{-1}(\mathbf{X}^*))\mathbf{c}^T\mathbf{X} \tag{14}$$

## III. Multivariate Markov Processes

In the Markovian approach we can impose a particular kind of dependence structure - conditional dependence - on $\mathbf{Y}(S^t)$, now considering a discrete time state space as in Vangas

The ordinary setup for a *continuous space Markov model* is to progress a process through

time; the Markovian property holds that this progression occur with a stable transition probability law.

For ease of notation let $s_1 = S^{t=1}$, etc. Then the Chapman-Kolmogorov equations

$$f_{\mathbf{Y}_{s_1},\ldots,\mathbf{Y}_{s_n}} = \int_{-\infty}^{\infty} f_{\mathbf{Y}_{s_n}|\mathbf{Y}_{s_{n-1}}}(\mathbf{Y}_{s_n}|\mathbf{Y}_{s_{n-1}}) \cdots f_{\mathbf{Y}_{s_2}|\mathbf{Y}_{s_1}}(\mathbf{Y}_{s_2}|\mathbf{Y}_{s_1}) d\mathbf{Y}_{s_2} \cdots d\mathbf{Y}_{s_{n-1}} \tag{15}$$

hold that the progression of the random process $\mathbf{Y}_{s_i}$ is governed by these transition probabilities, 'averaging' probability mass over the conditionally independent states.

### A. Copula approach

Take $Z_1 \sim F_{Z_1},\ Z_2 \sim F_{Z_2}$ (two components of $\mathbf{Z}$, say) and set $U = F_{Z_1}$ and $V = F_{Z_2}$; the pair $(U, V)$ are the '*grades*' of $(Z_1, Z_2)$ i.e. the mapping of $(Z_1, Z_2)$ in $F_{Z_1}, F_{Z_2}$ space. A copula is a function that takes the 'grades' as arguments and returns a joint distribution function, with marginals $F_{Z_1}, F_{Z_2}$.

$$C(U, V) = F_{Z_1,Z_2}$$

Any multivariate distribution function can yield a copula function,

$$F_{Z_1,Z_2}(F_{Z_1}^{-1}(U), F_{Z_2}^{-1}(V)) = C'(U, V)$$

that it: the correspondence which assigns the value of the joint distribution function to each ordered pair of values $(F_{Z_1}, F_{Z_2})$ for each $Z_1, Z_2$ is a distribution function called a copula (Nelsen 1996).

Joint distributions are specified by *marginal* and *dependence* parameters; for example a bivariate exponential distribution

$$H_\theta(x_1, x_2) = 1 - e^{-\lambda_1 x_1} - e^{-\lambda_2 x_2} + e^{-(\lambda_1 x_1 + \lambda_2 x_2 + \theta x_1 x_2)} \tag{16}$$

has marginal parameters $\lambda_1, \lambda_2$ and dependence parameter $\theta$. The copula version for this joint distribution is

$$C_\theta(u, v) = H(-ln(1-u), -ln(1-v)) =$$

$$= (u+v-1) + (1-u)(1-v) * e^{-\theta \ln(1-u) \ln(1-v)}$$

and the marginal parameters, still extant, are sublimated in the probability integral transformation of $U = F_{Z_1;\lambda_1}, V = F_{Z_2;\lambda_2}$

### B. Markov Processes via Copula: Darsow, Nguyen and Olsen

Following Darsow, Nguyen, Olsen (1991), define

$$(A * B)(Z_1, Z_2) = \int_0^1 \frac{\partial A(Z_1, t)}{\partial Z_2} \cdot \frac{\partial B(t, Z_2)}{\partial Z_1} dt$$

for $A, B$ copulas and $Z_1, Z_2$ in $I$. Since, for $Z_1, Z_2 \sim F_{Z_1}, F_{Z_2}, C$

$$\mathbb{P}(Z_1 < z_1 | Z_2 = z_2) = \frac{\partial C(F_{Z_1}, F_{Z_2})}{\partial Z_2}$$

and

$$\mathbb{P}(Z_2 < z_2 | Z_1 = z_1) = \frac{\partial C(F_{Z_1}, F_{Z_2})}{\partial Z_1}$$

then, for any three random variables $Z_1, Z_2, Z_3$, where $(Z_1 \perp Z_3)|Z_2$

$$C_{Z_1 Z_3} = C_{Z_1 Z_2} * C_{Z_2 Z_3}$$

Calling $C_{t_i t_j}$ the copula of the random variables $Z_{t_i}$, $Z_{t_j}$, then, for $t_i < t_j < t_k$

$$C_{t_i t_k} = C_{t_i t_j} * C_{t_j t_k} \tag{17}$$

is an equivalent representation of the CK equations, and

$$\mathbb{P}(Z_t \in A | Z_s = z) = \frac{\partial C_{st}(F_s(z), F_t(a))}{\partial Z_s} \tag{18}$$

is the copula version of the CK transition probability.

### *C. 'Tunable' Markovian models via Parametric Copula*

A markov process is 'conventionally' specified by a set of initial distributions $F_0$ and a family of transition probabilities $f_{Z_i|Z_j}(Z_i|Z_j)$; as an estimation problem, the goal is to estimate these transition probabilities from data.

In this copula based approach we assign the marginal distributions for each state $F_{Z_1}$,..., $F_{Z_m}$, and specify family of copulas satisfying (17). The estimation problem here is to fit the copulae, i.e. the *transition dependence* between states, from data. This is just to write (17) as

$$C_{t_i t_k;\theta_1,\theta_2} = C_{t_i t_j;\theta_2} * C_{t_j t_k;\theta_1}. \tag{19}$$

This yields a likelihood type method

$$(\hat{\theta}_1, \hat{\theta}_2) = \arg\max_{\theta_1,\theta_2} C_{t_i t_k;\theta_1,\theta_2} = C_{t_i t_j;\theta_2} * C_{t_j t_k;\theta_1}. \tag{20}$$

for fitting copula as transition probabilities, and an interpretation of the estimated parameters as the *transitional dependence* for the fitted Markov process. The copula dimensions match that of the transition probabilities: the simplest and special case is for 2-copula for pairwise conditional transitions.

This method is especially useful in Markov process estimation problems where: marginal distributions are available for each state; where the initial distribution of the process is non-informative; and where parametric models for *transition dependence* are desirable.

**Algorithm**

1) Initialize data and prior distributions
2) Compute dependent variables and prior distributions
3) Metropolis Hastings to compute posterior distributions
4) Step Markov Process – generate 'new' data from posterior
5) Compute dependent variables and prior distributions
6) Repeat until convergence in posterior distributions and reforested locations

**Fig. 1** Algorithm

Recall that each of these matrices are instantiated values of random variables in this setup. We consider each entry as independent, thus the distributions for each random matrix - at this moment - are concatenations of independent identically distributed random variables.

## IV. Bayesian Setup

The key to this approach is replacing the optimization scheme - i.e. the constrained objective function for Effective Accumulation - with the functional relationship between the independent and dependent data and then drawing from the probabilistic posterior of the control variables of interest: Effective Accumulation, the constraints on flow, and which nodes are to be reforested.

Each of these are now probabilistic with a posterior distribution; the choice of nodes to be reforested is a posterior probability for each node between zero and one.

The outline of the algorithm is listed in Figure 1. In fact the algorithm exploits Markovian theory twice: in step 3 we use the Metropolis-Hastings procedure to calculate the posterior distributions of the (transformed) data; in step 4 we recast the heuristic search in Vanegas [2009] as a Markov Chain, draw 'new' data iterates and return to step 2. Step 3 could be avoided in cases where conjugate posterior distributions have closed or analytic form.

We illustrate the algorithm using a low dimensional example, the same used in Vanegas [2009]. While we have some concern about the computational expensiveness of this procedure as we are able to parameterize the overarching Markov Chain: the analog for the heuristic search.

### *A. Data initialization*

We use the 'toy' data in Fig. 1(a) from Vanegas [2009]. These are five $4 \times 4$ matrices. The data are:

$$\mathcal{X}_{FD} = \begin{bmatrix} 5 & 5 & 5 & 0 \\ 4 & 4 & 4 & 3 \\ 4 & 5 & 3 & 3 \\ 4 & 3 & 3 & 2 \end{bmatrix} \quad (21)$$

$$\mathcal{X}_{FP,0} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (22)$$

$$\mathcal{X}_{FF,0} = \begin{bmatrix} 1 & .5 & .5 & 1 \\ 1 & .4 & .7 & .2 \\ 1 & .2 & .7 & .2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (23)$$

$$\mathcal{X}_{BP1,0} = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .5 & .5 & .5 & .5 \\ .5 & .5 & .5 & .5 \\ .5 & .5 & .5 & .5 \end{bmatrix} \quad (24)$$

and

$$\mathcal{X}_{BP2,0} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (25)$$

**Fig. 2** Matrix representation of data

So each of these are draws, singly valued replicates from independent identically distributed prior distributions for each variable.

*B. Compute Dependent Variables, Prior Distributions*

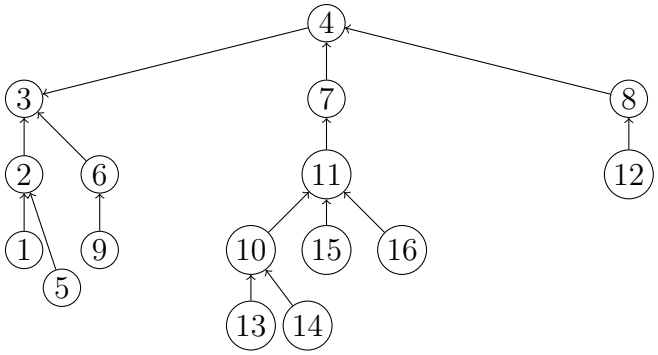


**Fig. 3** Graphical representation of data in Figure 1

*C. Metropolis-Hastings*

We use a modified version of the Metropolis-Hastings algorithm (Chib & Greenberg [1995]; Metropolis & Hastings [1953]) to draw replicates from the posterior distributions necessary for the third step in the Algorithm.

Here we exploit the fact that all of the transformations from 'data' (equations (1)-(5)) to the 'dependent variables' necessary for the calculation of Effective Accumulation are linear. The prior distributions in equations (1)-(5) are invariant to linear transforms - only the support must be adjusted; the distribution for equation (6) changes from Bernoulli to Binomial under linear transform.

In our set up then we use the prior distributions of the dependent variables (equation (14)) as **the** conditionally independent distributions we cycle through to generate posterior replicates from the joint multivariate distribution.

**Sub-Algorithm (2)-(3)**

1) Use Equation (14) to calculate prior distributions for $\mathbf{y} = (y_1, ..., y_K)$
2) Let $y_1[0] = y_1$, set $s = 0$
3) WHILE $s < S$ DO
4) FOR $k$ in $1 : K$
5) Choose $y_{k+1}[s+1]|y_k[1], ..., y_1[s]$ according to the density in Equation (14)
6) NEXT k
7) END WHILE

**Fig. 4** SubAlgorithm

We set $S$ sufficiently large to establish the ergodic properties of the resampled empirical distribution; for low dimensional data and these simple distributions $S$ on the order of $10^2$ should be sufficient, discarding the first several.

### *D. Markov Chain Progression*

Here we are replacing the heuristic search in Vanegas [2009] with a multivariate, discrete-time, real-valued Markov Chain.

In the original setup, the optimization is to minimize EA at the root node - of course dependent upon updated EA calculation at parent nodes. Once a node is reforested, EA is recalculated and downstream (or parent) nodes are considered for reforestation by recalculation of EA.

In our setup, we exploit a parametric representation of the *transition probability* between states in a Markov Chain (equation (18)). Looking again at this equation and modifying the notation:

$$\mathbb{P}(\mathbf{Z}_{t+1} \in A | \mathbf{Z}_t = \mathbf{z}) = \frac{\partial C^{\theta}_{t,t+1}(F_t(\mathbf{z}), F_{t+1}(A))}{\partial \mathbf{Z}_t} \tag{26}$$

Let $A \in \mathbf{A}$ be a particular state of reforestation for the nodes; if there are $K$ nodes (i.e. the dimension of each of the data matrices is $(\sqrt{K} \times \sqrt{K})$) then there are $2^K$ such states. Essentially the Vanegas heuristic search maps a subset of these possible states to an ordered list. In this notation $\mathbf{Z}_t$ is the complete set of information at discrete time $t$: the data matrices, the dependent variables, the set of nodes that are reforested at $t$, the

Our version is to impose a 'preferential' distribution on this multivariate transition probability; let $\theta \in \Theta$ be the particular choice for the parameterization for the copula $C$ governing the transitional dependency. Our job is only to select a particular choice for $\theta$ that 'matches' the optimization problem, in this case to minimize the EA at the root node.

Such a $\theta$ can be selected via Equation (20), using output from the Vanegas algorithm - or any set of reforested nodes with initial conditions - as data. Equation (20) generates maximum likelihood estimates (MLE) - $\hat{\theta}$ - for $\theta$. Then the particular Multivariate Markov Chain (out of a family of Markov Chains on $\Theta$, characterized by their transition probabilities of which a particular has been chosen by MLE $\hat{\theta}$)

**Sub-Algorithm (4)-(5)**

1) Choose $K$ dimensional copula $C$; Use Equation (20) to choose $\hat{\theta}$
2) Set $t = 0$
3) WHILE $t < T$ DO
4) Use Equation (26) to generate $\mathbf{Z}_{t+1}$ from $\mathbf{Z}_t$
5) GOTO SubAlgorithm (2)-(3), RETURN from
6) END WHILE

**Fig. 5** SubSubAlgorithm

In this way the Markov Chain will be updated by the deterministic (affine) transformations and the probabilistic draws from the Gibbs sampling in the Sub-Algorithm.

### *E. Semi-Fully Bayesian*

In notation we consider $\mathbf{Z}_t$ is the complete state of the problem at each discrete time unit. Via Equation (14) and Equation (20) we have the conditional (posterior) distribution for $\mathbf{Z}_t$ as

$$f_{\mathbf{Z}} \propto \frac{\partial C^{\theta}_{t,t+1}(F_t(\mathbf{z}), F_{t+1}(A))}{\partial \mathbf{Z}_t} \cdot f_{\mathbf{X}}(g^{-1}(\mathbf{X}))|\mathbf{J}| \tag{27}$$

where we have two crucial omissions. The first, the normalizing constant $g(\mathbf{Z})$, the marginal distribution for all the data, we access via the Gibbs sub algorithm. The second, a prior distribution for the parameter $\theta$, say $\pi(\theta)$, we consider non-informative in this problem. In future versions - i.e. where more data are available from an optimization - or better yet a real world re-forestation scheme - imposing an informative prior on the transitional probability function of the Markov Chain will be useful.

## V. Results

This approach, on low dimensional data, will not 'beat' the method in Vanegas [2009] in computational time; while we are replacing a a heuristic search with convergence on $O(n^2)$ (the heuristic search in Vanegas is essentially a sorting algorithm) with a probability estimation problem on $O(n^2)$, the additional draw from the posterior distribution introduces an additional $O(n^K)$ of computation.

Our goal here is more to demonstrate that the search can be placed in the parameter space of a properly determined probability model. We consider our measure successful if the results are similar to those of the heuristic search - which, are, as well, only estimates (but strictly deterministic ones) of a unique solution to an LP problem.

Using the 'toy' data from Vanegas, we initialize the matrices as random variables, use the affine probability transformations to propagate the network flow (effective accumulation) and the Markovian models via the Copulas to compute probabilities for Effective Accumulation in each node.

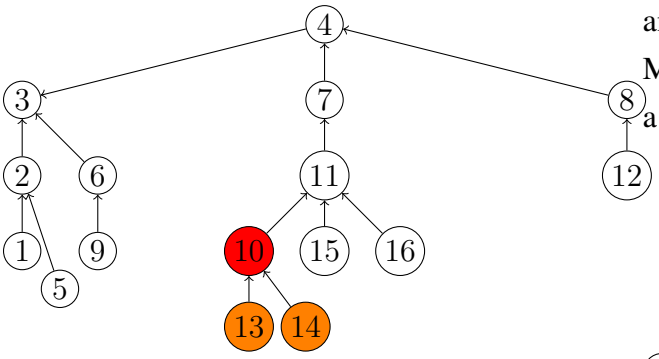


**Fig. 6** Graphical representation of data, one draw, first iteration

In the 'draw' illustrated in these figures - we use treat the first draw from the priors as degenerate: we just use the data as given in Vanegas [2009].

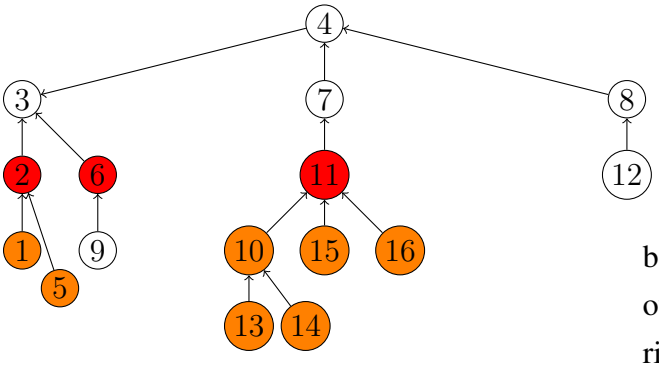


**Fig. 7** Graphical representation of data, one draw, second iteration

The nodes are colored in the figures as follows: no color - the mean of the posterior distribution for Effective Accumulation is lower than the mean for all nodes; orange - the mean of the posterior distribution for EA is between the mean and one standard deviation above it; red - the mean of the posterior for EA is above the mean plus one standard deviation. Remember that each 'iteration' is just a change in state (i.e. another draw from the posterior and updating of the affine equations) in the Multivariate Markovian process: each node has a posterior distribution for EA.

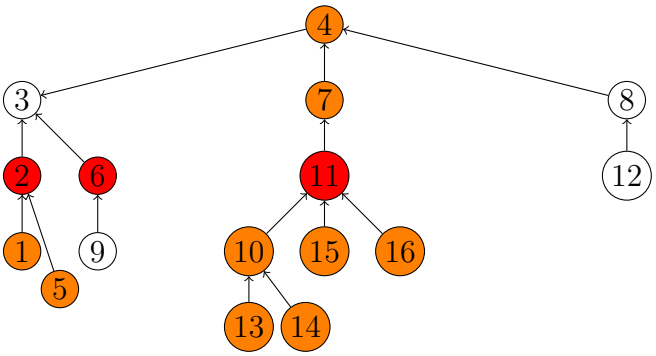


**Fig. 8** Graphical representation of data, one draw, third iteration

In practice these probabilities and rules can be defined with more granularity - for a model over many locations differences in these posterior probabilities may be slight.

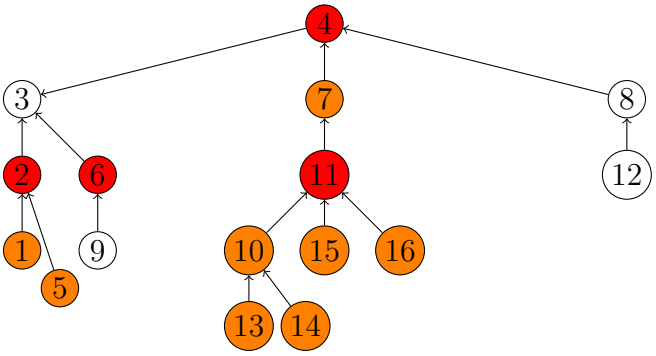


**Fig. 9** Graphical representation of data, one draw, fourth iteration

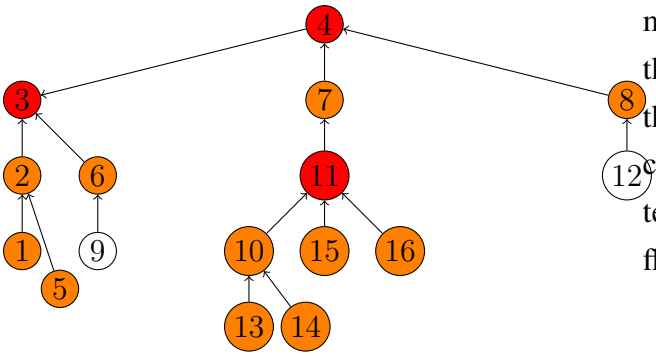


**Fig. 10** Graphical representation of data, one draw, fifth iteration

As well, for a high dimensional problem the computation time between iterates may be long. Not only do the affine equations need to be update, but the posterior distributions of the variates which depend on those equations.

Lastly - we chose an Archimedean copula as the 'engine' for this Multivariate Markov - other choices are popular. The Archimedean families have properties - including commutativity in arguments (abelian-ness) which may be undesirable in other modeling

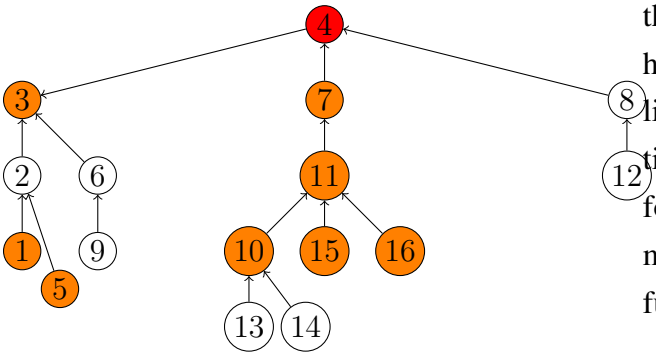


**Fig. 11** Graphical representation of data, one draw, sixth iteration

In this example we stopped running the algorithm after twenty-five iterations. Again - a larger dimension problem may require many more iterations, thus many more draws from the posterior distribution. The distribution of the multivariate process was 'stable' (no nodes changed color given our 'rules') after seventeen iterates though the probability values did fluctuate a bit.

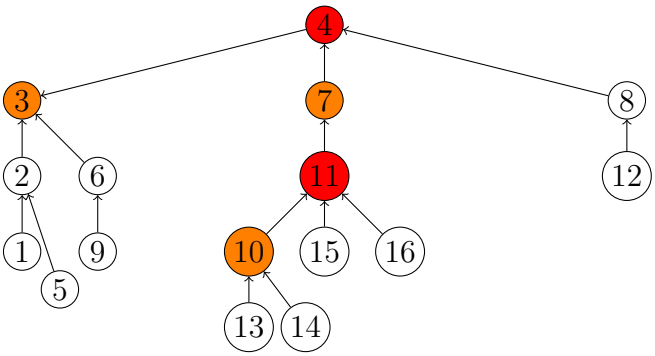


**Fig. 12** Graphical representation of data, one draw, twenty-fifth iteration

In summary: we have demonstrated how to replace a heuristic method to solve an LP problem with a probabilistic one. Mathematically, however, we have not demonstrated equivalence between the two, nor have we proven that our method *will* generate a solution for any heuristic method. We have illustrated that - on a linear control problem - Markovian probabilistic progression between states can substitute for deterministic equations. We address the mathematical support for this equivalence in future work.

## REFERENCES


Chib, Siddartha and Greenberg, Edward (1995) "Understanding the Metropolis-Hastings Algorithm" *The American Statistician*. pp. 327-335,

Darsow, W., Nguyen, B, and Olsen, E. "Copulas and markov processes." *Illinois Journal of Mathematics*, 36:600

Hastings, W.K. (1970): "Monte Carlo Sampling Methods Using Markov Chains and Their Applications" *Biometrika*, 57, 97-109.

Koller, D., Friedman, N. *Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning Series)*. MIT Press. 2009

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller (1953): "Equations of State Calculations by Fast Computing Machines" *Journal of Chemical Physics*, 21, 1087-1092.

Vanegas, P, et al "Allocating Reforestation Areas for Sediment Flow Minimization: An Integer Programming Formulation and a Heuristic Solution Method." Manuscript. 2009.